

PEACE PUBLISHERS



Н. В. ЕФИМОВ

КРАТКИЙ КУРС
АНАЛИТИЧЕСКОЙ ГЕОМЕТРИИ

ГОСУДАРСТВЕННОЕ ИЗДАТЕЛЬСТВО
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ

Москва

N. YEFIMOV

A Brief Course
in
Analytic Geometry

TRANSLATED FROM THE RUSSIAN
by O. SOROKA

PEACE PUBLISHERS
MOSCOW

Written by Professor N. Yefimov, Dr. Phys. Math. Sc., this book presents, in concise form, the theoretical foundations of plane and solid analytic geometry. Also, an elementary outline of the theory of determinants is given in the Appendix.

The textbook is intended for students of higher educational institutions and for engineers engaged in the field of quadric surface design.

The book is illustrated with 122 drawings.

516.3
7351

CONTENTS

Part One

PLANE ANALYTIC GEOMETRY

<i>Chapter 1.</i> Coordinates on a Straight Line and in a Plane	11
§ 1. An Axis and Segments of an Axis	11
§ 2. Coordinates on a Line. The Number Axis	14
§ 3. Rectangular Cartesian Coordinates in a Plane. A Note on Oblique Cartesian Coordinates	16
§ 4. Polar Coordinates	20
 <i>Chapter 2.</i> Elementary Problems of Plane Analytic Geometry	 23
§ 5. Projection of a Line Segment. Distance Between Two Points	23
§ 6. Calculation of the Area of a Triangle	28
§ 7. Division of a Line Segment in a Given Ratio	30
§ 8. Transformation of Cartesian Coordinates by Translation of Axes	35
§ 9. Transformation of Rectangular Cartesian Coordinates by Rotation of Axes	36
§ 10. Transformation of Rectangular Cartesian Coordinates by Change of Origin and Rotation of Axes	38
 <i>Chapter 3.</i> The Equation of a Curve	 41
§ 11. The Concept of the Equation of a Curve. Examples of Curves Represented by Equations	41
§ 12. Examples of Deriving the Equation of a Given Curve	48
§ 13. The Problem of the Intersection of Two Curves	50
§ 14. Parametric Equations of a Curve	51
§ 15. Algebraic Curves	52
 <i>Chapter 4.</i> Curves of the First Order	 55
§ 16. The Slope of a Straight Line	55
§ 17. The Slope-intercept Equation of a Straight Line	56
§ 18. Calculation of the Angle Between Two Straight Lines. Conditions for the Parallelism and Perpendicularity of Two Straight Lines	59
§ 19. The Straight Line As the Curve of the First Order. The General Equation of the Straight Line	61
§ 20. Incomplete Equations of the First Degree. The Intercept Equation of a Straight Line	63
§ 21. Discussion of a System of Equations Representing Two Straight Lines	65
§ 22. The Normal Equation of a Straight Line. The Problem of Calculating the Distance of a Point from a Straight Line	68
§ 23. The Equation of a Pencil of Lines	71
 <i>Chapter 5.</i> Geometric Properties of Curves of the Second Order	 75
§ 24. The Ellipse. Definition of the Ellipse and Derivation of Its Canonical Equation	75

§ 25. Discussion of the Shape of the Ellipse	79
§ 26. The Eccentricity of the Ellipse	81
§ 27. Rational Expressions for Focal Radii of the Ellipse	82
§ 28. Point-by-point Construction of the Ellipse. The Parametric Equations of the Ellipse	83
§ 29. The Ellipse as the Projection of a Circle on a Plane. The Ellipse as the Section of a Circular Cylinder by a Plane	84
§ 30. The Hyperbola. Definition of the Hyperbola and Derivation of Its Canonical Equation	87
§ 31. Discussion of the Shape of the Hyperbola	91
§ 32. The Eccentricity of the Hyperbola	97
§ 33. Rational Expressions for Focal Radii of the Hyperbola	98
§ 34. The Directrices of the Ellipse and Hyperbola	99
§ 35. The Parabola. Derivation of the Canonical Equation of the Parabola	103
§ 36. Discussion of the Shape of the Parabola	105
§ 37. The Polar Equation of the Ellipse, Hyperbola and Parabola	107
§ 38. Diameters of Curves of the Second Order	109
§ 39. The Optical Properties of the Ellipse, Hyperbola and Parabola	114
§ 40. The Ellipse, Hyperbola and Parabola as Conic Sections	115
<i>Chapter 6. Transformation of Equations by Change of Coordinates</i>	<i>116</i>
§ 41. Examples of Reducing the General Equation of a Second-Order Curve to Canonical Form	116
§ 42. The Hyperbola as the Inverse Proportionality Graph. The Parabola as the Graph of a Quadratic Function	125
 Part Two SOLID ANALYTIC GEOMETRY 	
<i>Chapter 7. Some Elementary Problems of Solid Analytic Geometry</i>	<i>131</i>
§ 43. Rectangular Cartesian Coordinates in Space	131
§ 44. The Concept of a Free Vector. The Projection of a Vector on an Axis	134
§ 45. The Projections of a Vector on the Coordinate Axes	138
§ 46. Direction Cosines	141
§ 47. Distance between Two Points. Division of a Line Segment in a Given Ratio	142
<i>Chapter 8. Linear Operations on Vectors</i>	<i>143</i>
§ 48. Definitions of Linear Operations	143
§ 49. Basic Properties of Linear Operations	144
§ 50. The Vector Difference	147
§ 51. Fundamental Theorems on Projections	149
§ 52. Resolution of Vectors into Components	152
<i>Chapter 9. The Scalar Product of Vectors</i>	<i>157</i>
§ 53. The Scalar Product and Its Basic Properties	157
§ 54. Representation of the Scalar Product in Terms of the Coordinates of the Vector Factors	160
<i>Chapter 10. The Vector and Triple Scalar Products of Vectors</i>	<i>163</i>
§ 55. The Vector Product and Its Basic Properties	163
§ 56. Representation of the Vector Product in Terms of the Coordinates of the Vector Factors	170

§ 57. The Triple Scalar Product	172
§ 58. Representation of the Triple Scalar Product in Terms of the Coordinates of the Vector Factors	176
<i>Chapter 11. The Equation of a Surface and the Equations of a Curve . . .</i>	178
§ 59. The Equation of a Surface	178
§ 60. The Equations of a Curve. The Problem of the Intersection of Three Surfaces	179
§ 61. The Equation of a Cylindrical Surface with Elements Parallel to a Coordinate Axis	181
§ 62. Algebraic Surfaces	183
<i>Chapter 12. The Plane as the Surface of the First Order. The Equations of a Straight Line</i>	185
§ 63. The Plane as the Surface of the First Order	185
§ 64. Incomplete Equations of Planes. The Intercept Form of the Equa- tion of a Plane	188
§ 65. The Normal Equation of a Plane. The Distance of a Point from a Plane	190
§ 66. The Equations of a Straight Line	194
§ 67. The Direction Vector of a Straight Line. The Canonical Equations of a Straight Line. The Parametric Equations of a Straight Line .	198
§ 68. Some Additional Propositions and Examples	201
<i>Chapter 13. Quadric Surfaces</i>	207
§ 69. The Ellipsoid and the Hyperboloids	207
§ 70. The Quadric Cone	213
§ 71. The Paraboloids	215
§ 72. The Quadric Cylinders	219
§ 73. The Rectilinear Generators of the Hyperboloid of One Sheet. The Shukhov Towers	220
<i>Appendix. The Elements of the Theory of Determinants</i>	225
§ 1. Determinants of the Second Order and Systems of Two Equations of the First Degree in Two Unknowns	225
§ 2. A Homogeneous System of Two Equations of the First Degree in Three Unknowns	229
§ 3. Determinants of the Third Order	232
§ 4. Cofactors and Minors	236
§ 5. Solution and Analysis of a System of Three First-degree Equa- tions in Three Unknowns	240
§ 6. The Concept of a Determinant of Any Order	247

PART ONE

**Plane Analytic
Geometry**

Chapter I

COORDINATES ON A STRAIGHT LINE AND IN A PLANE

§ 1. An Axis and Segments of an Axis

1. Consider an arbitrary straight line. It extends in two opposite directions. Let us choose at will one of these directions and refer to it as positive (and to the other direction as negative).

A straight line to which a positive direction has been assigned is called *an axis*. In diagrams, the positive direction of an axis is indicated by an arrowhead (see, for example, Fig. 1, where an axis a is shown).

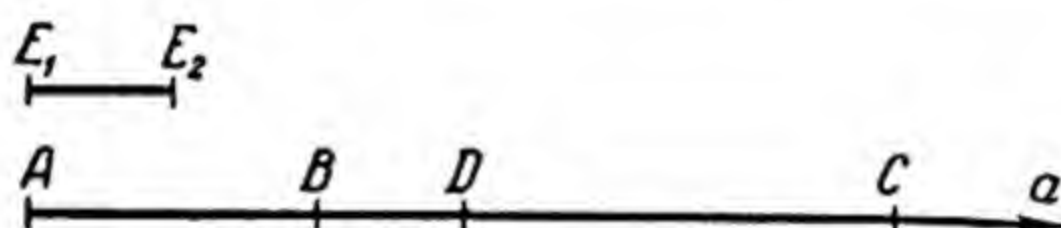


Fig. 1.

2. Let there be given an axis, and let there also be specified a *unit segment*, that is, a unit of length by means of which any segment can be measured, so that the length of any segment may be determined.

Take two arbitrary points on a given axis and denote them by the letters A and B . The segment bounded by the points A , B is called a *directed segment* if one of these points has been designated as the *initial point*, and the other as the *terminal point* of the segment. A segment is considered to be directed from its initial to its terminal point.

In the succeeding pages, a directed segment is denoted by the two letters that mark the points bounding the segment; the letter marking the initial point is always written first, and a horizontal bar is placed over the letters. Thus, \overline{AB} denotes the directed segment bounded by the points A , B and having A as its initial point; \overline{BA} denotes the directed segment bounded by the points A , B and having B as its initial point.

When dealing with directed segments of an axis in our future work, we shall often refer to them simply as segments, omitting the word "directed".

Let us now agree to define the value of a segment \overline{AB} of an axis as the number equal to the length of the segment and taken with a plus or minus sign according as the direction of the segment agrees with the positive or the negative direction of the axis. We shall denote the value of a segment \overline{AB} by the symbol AB (without the horizontal bar). We do not exclude the case when the points A and B coincide; in this case, the segment \overline{AB} is called a *zero segment*, since its value is equal to zero. A zero segment has no definite direction, and hence the term "directed" may be applied to such a segment only conventionally.

The *value* of a segment, as distinct from its *length*, is a *signed* number; the length of a segment is obviously the modulus *) of its value, and therefore, in accordance with the notation adopted in algebra for the modulus of a number, we shall use the symbol $|AB|$ to denote the length of a segment \overline{AB} . Clearly, $|AB|$ and $|BA|$ represent the same number. On the other hand, the values AB and BA themselves differ in sign, so that

$$AB = -BA.$$

Figure 1 shows an axis a and points A, B, C, D on that axis; E_1E_2 is a unit segment. The points A, B, C, D are assumed to occupy positions such that the distance between A and B is equal to two units, and the distance between C and D to three units. The direction from A to B agrees with, and the direction from C to D is opposite to, the positive direction of the axis. Accordingly, we have

$$AB = 2, \quad CD = -3,$$

or

$$BA = -2, \quad DC = 3.$$

Also, we may write

$$|AB| = 2, \quad |CD| = 3.$$

3. For any position of points A, B, C on an axis, the values of the segments \overline{AB} , \overline{BC} and \overline{AC} are connected by the relation

$$AB + BC = AC; \tag{1}$$

we shall call this relation the *fundamental identity*.

Let us prove this fundamental identity. Suppose first that the segments \overline{AB} and \overline{BC} are different from zero and have the *same* direction (Fig. 2, the top line); then the segment \overline{AC} has its

*) The word "modulus" has the same meaning as "absolute value".

length equal to the sum of the lengths of the segments \overline{AB} , \overline{BC} and agrees in direction with these segments. In this case, all the three numbers AB , BC and AC have like signs, and the number AC is equal to the sum of the numbers AB , BC , which means that identity (1) holds true.

Suppose next that the segments \overline{AB} and \overline{BC} are different from zero and have *opposite* directions (Fig. 2, the bottom line). Then the segment \overline{AC} has its length equal to the difference of the lengths of the segments \overline{AB} , \overline{BC} and agrees in direction with the longer of these segments. In

this case, the numbers AB and BC differ in sign, and AC has its modulus equal to the difference of the moduli of the numbers AB , BC and agrees in sign with the num-

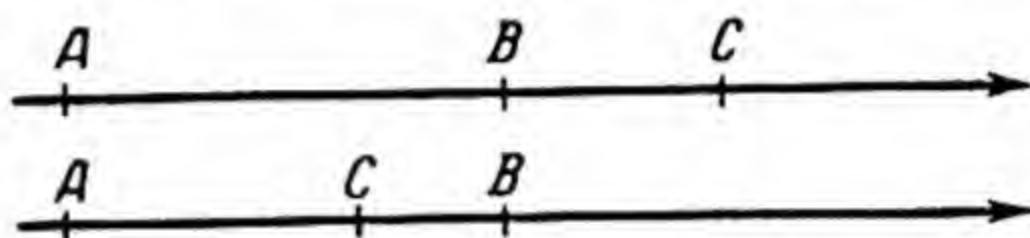


Fig. 2.

ber having the larger modulus. Consequently, for this position of the points, the number AC is equal to the sum of the numbers AB , BC , according to the rule for addition of signed numbers; this means that identity (1) is again valid.

Finally, suppose that one of the segments \overline{AB} , \overline{BC} is a zero segment. If \overline{AB} is a zero segment, then the point B coincides with the point A , and hence

$$AB + BC = AA + AC = 0 + AC = AC.$$

If \overline{BC} is a zero segment, then the point B coincides with the point C , and hence

$$AB + BC = AC + CC = AC + 0 = AC.$$

Thus, identity (1) is actually valid for every position of the points A , B , C .

Note. If the symbols AB , BC and AC in relation (1) were considered simply as the *lengths* of the respective segments (without regard to sign!), the relation would be valid only for the case when the point B is situated between the points A and C . Relation (1) owes its universality precisely to the fact that AB , BC and AC are understood in it as the *values* of the segments \overline{AB} , \overline{BC} and \overline{AC} , that is, as their lengths taken with appropriate signs *).

*) In the case of segments not lying on an axis and regarded as arbitrary segments in a plane, no sign need be attributed to their lengths. The lengths of such segments may be denoted as in elementary geometry (that is, without the modulus symbol); this will often be done below (see, for example, Art. 40, where the length of the segment is denoted by CM , rather than by $|CM|$).

§ 2. Coordinates on a Line. The Number Axis

4. In this article we shall describe a method which permits us to determine the *position of points* on an arbitrarily chosen straight line by the specification of *numbers*.

Let there be given an arbitrary straight line a . We next choose a segment as the unit of length, assign a positive direction to the line a (thereby making it an axis), and mark some point on this line by the letter O .

Let us now agree to define the coordinate of any point M on the axis a as the value of the segment \overline{OM} . The point O will be called the origin of coordinates; the coordinate of the origin itself is equal to zero.

The specification of the coordinate of a point M completely determines the position of M on the given line. For, the modulus of the coordinate, i. e. $|OM|$, is the distance of M from the (pre-assigned) point O , while the sign of the coordinate, i. e., the sign of the number OM , gives the direction in which the point M lies relative to the point O ; if the coordinate is positive, then M lies in the positive direction from O ; if the coordinate is negative, then M lies in the negative direction from O ; if the coordinate is zero, then M coincides with O . (All this immediately follows from the definition of the value of a segment of an axis; see Art. 2.)

Imagine that the straight line a is drawn horizontally before us and is positively directed to the right. The location of points of the line a , according to the sign of their coordinates, may then be described as follows: Points with positive coordinates lie to the right of the origin O , and points with negative coordinates, to the left.

The coordinate of an arbitrary point is generally denoted by the letter x . In cases when several points are considered, they are often denoted by one letter having different subscripts, say M_1, M_2, \dots, M_n ; then the coordinates of these points are also denoted by one letter with corresponding subscripts: x_1, x_2, \dots, x_n .

To show briefly that a given point has a given coordinate, the latter is enclosed in parentheses and written next to the symbol of the point itself, as for example, $M_1(x_1), M_2(x_2), \dots, M_n(x_n)$.

5. We shall now prove two simple but important theorems concerning an axis to which a coordinate system has been attached.

Theorem 1. For any two points $M_1(x_1)$ and $M_2(x_2)$ of an axis, we always have the relation

$$M_1M_2 = x_2 - x_1. \quad (1)$$

Proof. In consequence of the fundamental identity (Art. 3),

$$OM_1 + M_1M_2 = OM_2,$$

whence

$$M_1M_2 = OM_2 - OM_1.$$

But $OM_2 = x_2$, $OM_1 = x_1$, and so

$$M_1M_2 = x_2 - x_1,$$

as was to be proved.

This theorem may be phrased as follows: *To find the value of a segment of an axis, it is necessary to subtract the coordinate of its initial point from the coordinate of its terminal point.* (See

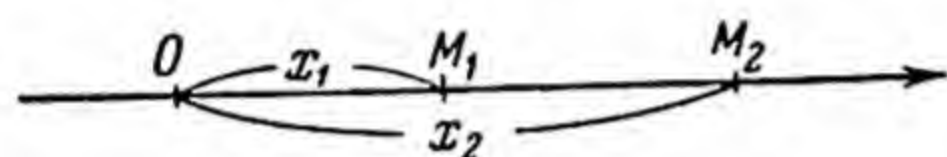


Fig. 3.



Fig. 4.

Figs 3 and 4; it should be observed that the coordinate x_1 in Fig. 4 is negative.)

Theorem 2. *If $M_1(x_1)$ and $M_2(x_2)$ are any two points of an axis, and if d is the distance between them, then*

$$d = |x_2 - x_1|. \quad (2)$$

Proof. In agreement with the preceding theorem,

$$M_1M_2 = x_2 - x_1;$$

but the distance between the points M_1 and M_2 is the modulus of the value of the segment $\overline{M_1M_2}$, and hence

$$d = |x_2 - x_1|.$$

The theorem is thus proved.

Note. Since the numbers $x_2 - x_1$ and $x_1 - x_2$ have the same modulus, it is equally correct to write $d = |x_2 - x_1|$ and $d = |x_1 - x_2|$. Taking this into account, we may phrase the above theorem thus: *To compute the distance between two points of an axis, it is necessary to subtract the coordinate of one point from the coordinate of the other and take the modulus of the resulting difference.*

Example 1. Given the points $A(5)$, $B(-1)$, $C(-8)$, $D(2)$, find the values of the segments \overline{AB} , \overline{CD} and \overline{DB} .

Solution. By Theorem 1, we have

$$AB = -1 - 5 = -6,$$

$$CD = 2 - (-8) = 10,$$

$$DB = -1 - 2 = -3.$$

Example 2. Find the distance between the points $P(3)$ and $Q(-2)$.

Solution. It follows from Theorem 2 that

$$d = |-2 - 3| = |-5| = 5.$$

6. If a coordinate system is attached to an axis, each point of that axis will have one completely determined coordinate. Conversely, for any (real) number x there will be found on the axis one completely determined point M having the given coordinate x .

Let us adopt the convention that the point M represents the number x . An axis on which coordinates have been introduced (by the method described in Art. 4) is called the *number axis*. Fig. 5. shows the number axis with several integers marked on it.

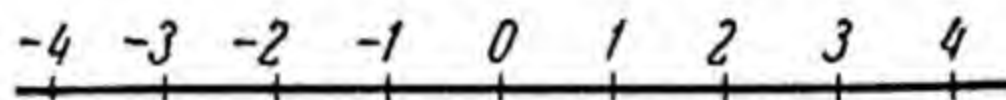


Fig. 5.

Representation of numbers as points of the number axis permits us to visualise geometrically our concept of the totality of numbers, and also makes it possible to express arithmetical relations in geometrical terms. For example, all solutions of the inequalities $3 < x < 5$ may be visualised as the points situated on the number axis between the point representing the number 3 (that is, having the coordinate 3) and the point representing the number 5 (that is, having the coordinate 5). This fact can be expressed briefly as follows: The inequalities $3 < x < 5$ represent the interval (on the number axis) bounded by the points 3 and 5.

The geometric representation of arithmetical relations has proved to be of great convenience and is widely used in all branches of mathematics.

§ 3. Rectangular Cartesian Coordinates in a Plane.

A Note on Oblique Cartesian Coordinates

7. When a method has been indicated, which permits us to establish the location of points in a plane by the specification of numbers, then we say that a *coordinate system* has been attached to the plane. We shall now consider the simplest and most com-

monly used coordinate system, called the *rectangular cartesian system of coordinates*.

A *rectangular cartesian system of coordinates* is determined by the choice of a linear unit (for measurement of lengths) and of two mutually perpendicular axes numbered (that is, designated as the first and second axis) in any order. The point of intersection of the axes is called the *origin of coordinates*, and the axes themselves are called the *coordinate axes*; the first axis is also termed the *x-axis* or *axis of abscissas*, and the second axis is termed the *y-axis* or *axis of ordinates*.

Let us denote the origin by the letter O , the x -axis by the letters Ox , and the y -axis by Oy . In diagrams, the letters x , y mark the respective axes at the points farthestmost from O in the positive direction, so that the direction of the x - and y -axes is unambiguously indicated by the position of the letters O , x and O , y , respectively. Thus, there is no need to indicate the positive directions of the coordinate axes by means of arrowheads, which will therefore be omitted in the succeeding diagrams.

Let M be an arbitrary point in the plane. Project the point M on the coordinate axes, that is, drop perpendiculars from M to the lines Ox and Oy ; mark the respective feet of these perpendiculars as M_x and M_y (Fig. 6).

The coordinates of a point M in a given system are defined as the numbers

$$x = OM_x, \quad y = OM_y, \quad (1)$$

where OM_x is the value of the segment $\overline{OM_x}$ of the x -axis, and OM_y is the value of the segment $\overline{OM_y}$ of the y -axis. The number x is called the *first coordinate* or *abscissa* of the point M , and the number y is called the *second coordinate* or *ordinate* of M . To indicate briefly that the point M has the abscissa x and the ordinate y , we use the notation $M(x, y)$. If several points are to be considered, we shall often denote them by one letter having different subscripts, as for example, M_1, M_2, \dots, M_n ; the coordinates of these points will then have corresponding subscripts, and so the points under consideration will be written thus: $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_n(x_n, y_n)$.

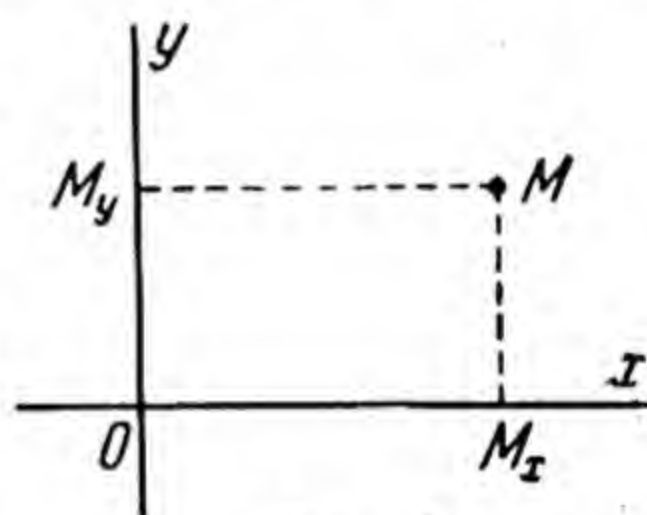


Fig. 6.

8. When a rectangular cartesian system of coordinates has been attached to a plane, then each point of the plane has one completely determined pair of coordinates x, y in this system. Conversely, for any two (real) numbers x, y , there will be found in the plane one completely determined point whose abscissa is x

and whose ordinate is y in the given system. To plot a point from its coordinates x, y , the segment $\overline{OM_x}$ equal in value to x is laid off (from the origin) on the x -axis, and the segment $\overline{OM_y}$ of the value y is laid off on the y -axis, the directions in which these segments are laid off being determined by the signs of the numbers x, y . Next, a line is drawn through M_x parallel to the axis Oy , and another line through M_y parallel to Ox ; the intersection of these two lines will give the required point M .

9. In Art. 4 it was explained how a system of coordinates is attached to a straight line. We shall now attach a system of coordinates to each of the coordinate axes Ox and Oy , retaining the given linear unit and the given directions of the axes Ox, Oy , and choosing the point O as the origin of coordinates on each axis.

Consider an arbitrary point M and its projection M_x on the axis Ox .

The coordinate of M_x on the axis Ox is equal to the value of the segment $\overline{OM_x}$; in Art. 7, this value was called the abscissa of the point M . Hence we conclude that the *abscissa of the point M is equal to the coordinate of the point M_x on the axis Ox* . Similarly, the *ordinate of the point M is equal to the coordinate of the point M_y on the axis Oy* . For all their obviousness, these propositions are of great importance; in fact, they enable us to apply Theorems 1 and 2 (Art. 5), which express basic properties of a coordinate system on a line, to all points in a plane.

10. To facilitate subsequent formulations, let us now agree about the use of certain terms.

The axis Oy divides the entire plane into two half-planes; the half-plane containing the positive half of the axis Ox will be termed *the right half-plane*, and the other, *the left half-plane*.

In like manner, the axis Ox divides the plane into two half-planes, of which the one containing the positive half of the axis Oy will be termed *the upper half-plane*, and the other, *the lower half-plane* *).

11. Let M be an arbitrary point of the right half-plane; then the segment $\overline{OM_x}$ is positively directed on the axis Ox and, consequently, the abscissa $x = \overline{OM_x}$ of the point M is positive. On the other hand, if M lies in the left half-plane, then the segment $\overline{OM_x}$ is negatively directed on the axis Ox , and the number $x = \overline{OM_x}$ is negative. Finally, if the point M lies on the axis Oy ,

*) The adoption of these terms is justified by the usual position of the coordinate axes in diagrams, where the axis Ox is generally seen directed to the right, and the axis Oy directed upwards.

its projection M_x on the axis Ox coincides with the point O and $x = OM_x$ is zero.

Thus, *all points of the right half-plane have positive abscissas ($x > 0$); all points of the left half-plane have negative abscissas ($x < 0$); the abscissas of all points lying on the axis Oy are equal to zero ($x = 0$).*

By similar reasoning, it can be established that *all points of the upper half-plane have positive ordinates ($y > 0$); all points of the lower half-plane have negative ordinates ($y < 0$); the ordinates of all points lying on the axis Ox are equal to zero ($y = 0$).*

Note that the origin O , as the point of intersection of the axes, has both coordinates equal to zero: $x = 0$, $y = 0$, and is characterised by this property (that is, both coordinates are zero *only* for the point O).

12. The two coordinate axes jointly divide the plane into four parts, called *quadrants* and numbered according to the following rule: The first quadrant is the one lying simultaneously in the right and the upper half-planes; the second quadrant lies in the left and the upper half-planes; the third quadrant lies in the left and the lower half-planes; and, finally, the fourth quadrant is the one situated in the right and the lower half-planes. (The order in which the quadrants are numbered is illustrated in Fig. 7.)

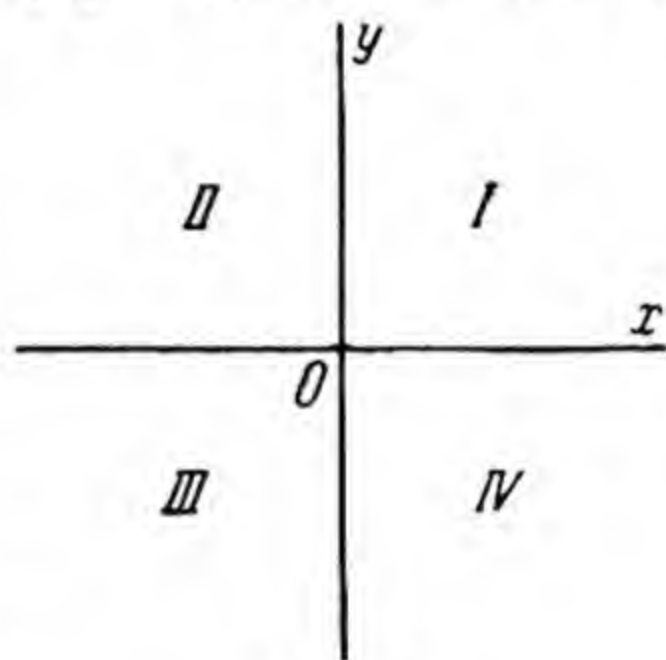


Fig. 7.

Let M be a point with coordinates x , y . From the foregoing, it follows that,

- if $x > 0$, $y > 0$, M lies in the first quadrant;
- if $x < 0$, $y > 0$, M lies in the second quadrant;
- if $x < 0$, $y < 0$, M lies in the third quadrant;
- if $x > 0$, $y < 0$, M lies in the fourth quadrant.

Consideration of the coordinate half-planes and quadrants is useful because it permits an easy orientation as to the position of the given points by the signs of their coordinates.

13. We have now become acquainted with the rectangular coordinate system, which is the most commonly used system of coordinates. However, when dealing with certain special problems, other systems may turn out to be more convenient in some cases. Let us therefore devote a few words to a cartesian coordinate system whose axes intersect at any angle.

Such a system is determined by choosing a unit of length and two axes Ox , Oy , which intersect in the point O at any angle (other than 0° and 180°). Let M be an arbitrary point of the plane. Draw two lines through M parallel to the axes Ox , Oy and denote their respective points of intersection with Ox and Oy by M_x and M_y (Fig. 8).

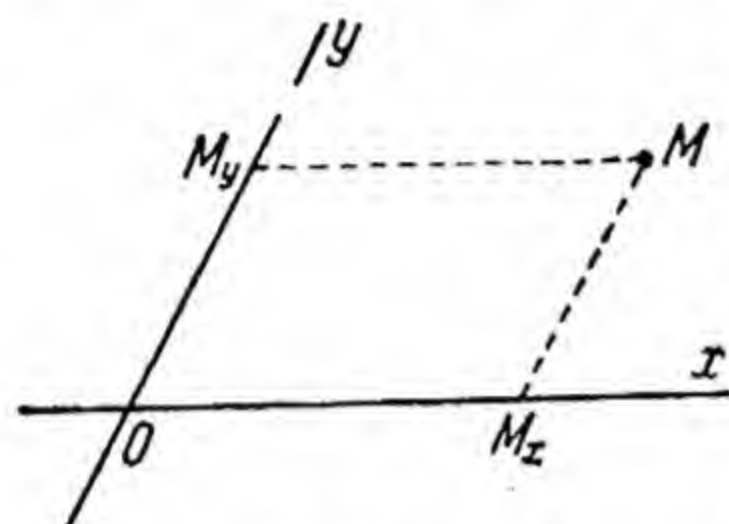


Fig. 8.

In the chosen system, the coordinates of the point M are defined as the numbers

$$x = OM_x, \quad y = OM_y,$$

where OM_x is the value of the segment $\overline{OM_x}$ of the axis Ox , and OM_y is the value of the segment $\overline{OM_y}$ of the axis Oy .

The *rectangular* cartesian system of coordinates constitutes the particular case of the above-described system when the axes Ox , Oy make a *right* angle. If the angle between the axes Ox , Oy is other than right, the system is called an *oblique* cartesian system of coordinates. Since oblique cartesian coordinates find no further application in this book, we shall often refer to *rectangular* cartesian coordinates simply as cartesian coordinates.

§ 4. Polar Coordinates

14. We shall now describe the so-called *polar coordinate system*, which is a very convenient and fairly frequently used system.

A polar coordinate system is determined by choosing a point O , called the *pole*, a ray OA , drawn from that point and called the *polar axis*, and a scale for measurement of lengths. When determining a polar system, it must also be specified which direction of rotation about the point O is to be considered positive. Usually, counter-clockwise rotation is taken as positive.

Given the pole and the polar axis (Fig. 9), let us consider an arbitrary point M . Let ρ denote the distance of M from the point O ($\rho = |OM|$), and θ the angle through which the ray OA must be rotated to reach coincidence with the ray OM ($\theta = \angle AOM$). This angle θ will be understood as in trigonometry (that is, as a signed quantity determined to within an addend of the form $\pm 2n\pi$).

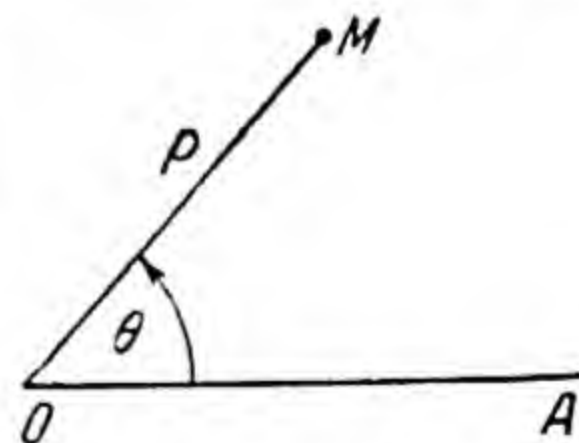


Fig. 9.

The numbers ρ and θ are called the *polar coordinates* of the point M (in reference to the chosen system). The number ρ is termed the *first coordinate* or *polar radius*, and the number θ , the *second coordinate* or *polar angle* (the polar angle is also termed the *amplitude*).

Note 1. From all possible values of the polar angle of the point M , one definite value is singled out, namely, the value which satisfies the inequalities

$$-\pi < \theta \leq \pi;$$

we shall call it the *principal value*. The principal value of the polar angle may be defined verbally as the angle through which the ray OA must be rotated (no matter which direction) to reach coincidence with the ray OM , so that the rotation will not exceed 180° . In the particular case when the direction of the ray OM is precisely opposite to that of the ray OA , two 180° rotations are possible; then the positive rotation is chosen, that is, $\theta = \pi$ is taken as the principal value of the polar angle.

Note 2. If the point M coincides with O , then $\rho = |OM| = 0$. This means that the first coordinate of the pole is equal to zero; its second coordinate obviously has no definite value.

15. Sometimes, the cartesian and the polar systems are to be used side by side. The following problem arises in such cases: Given the *polar* coordinates of a point, to find its *cartesian* coordinates; and, conversely, given the *cartesian* coordinates of a point, to find its *polar* coordinates. We shall now derive the formulas of this coordinate transformation (that is, the *formulas for transition from polar to rectangular coordinates, and vice versa*) for the particular case when the pole of the polar system coincides with the origin of rectangular cartesian coordinates, and the polar axis coincides with the positive semi-axis of abscissas (Fig 10). Also, when determining the polar angle, we shall regard as positive the direction of the shortest rotation of the positive semi-axis Ox into the positive semi-axis Oy .

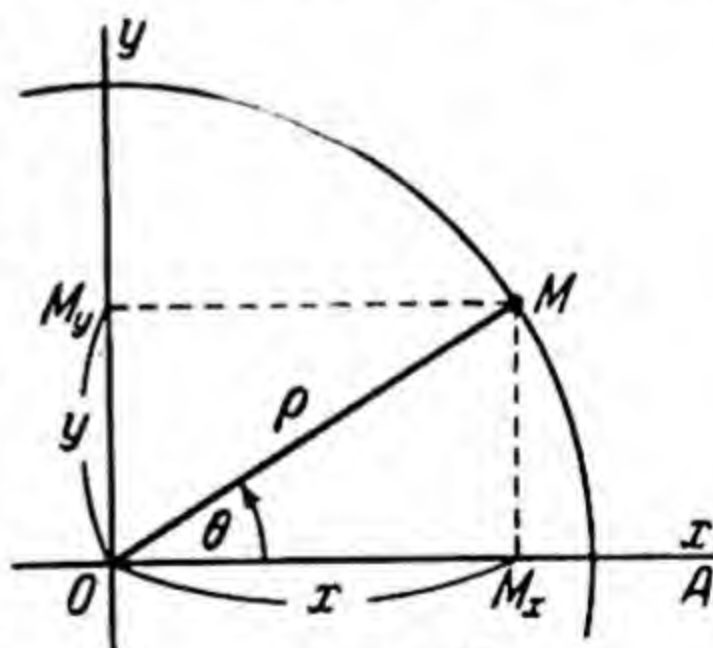


Fig. 10.

Let M be an arbitrary point in the plane, and let (x, y) and (ρ, θ) be its cartesian and polar coordinates, respectively. Describe a circle of radius ρ about the pole O ; this circle will be considered

as the trigonometric unit circle, and the axis Ox as the initial diameter. From the point M , drop perpendiculars to the axes Ox and Oy ; denote their feet by M_x and M_y respectively (see Fig. 10). The segment $\overline{OM_x}$ is the cosine line of the angle θ ; therefore, $OM_x = |OM| \cos \theta$. The segment $\overline{OM_y}$ is the sine line of the angle θ ; therefore $OM_y = |OM| \sin \theta$. But $OM_x = x$, $OM_y = y$, $|OM| = \rho$; hence, the above relations give

$$x = \rho \cos \theta, \quad y = \rho \sin \theta. \quad (1)$$

These are the formulas expressing cartesian coordinates in terms of polar coordinates. The expressions for polar coordinates in terms of cartesian coordinates may be obtained either from these formulas or directly from Fig. 10:

$$\rho = \sqrt{x^2 + y^2}, \quad \tan \theta = \frac{y}{x}. \quad (2)$$

Note, however, that the formula $\tan \theta = \frac{y}{x}$ fails to completely determine even the principal value of the polar angle; in fact, we must also know whether θ is positive or negative.

Example. Given the rectangular cartesian coordinates of a point: $(-2, 2)$; find its polar coordinates (assuming that the pole of the polar system coincides with the origin of the cartesian system, and that the polar axis coincides with the positive x -axis).

Solution. By formulas (2), we have

$$\rho = 2\sqrt{2}, \quad \tan \theta = -1.$$

According to the second of these relations, $\theta = \frac{3}{4}\pi$ or $\theta = -\frac{1}{4}\pi$. Since the given point lies in the second quadrant, the first of the two indicated values of θ must be chosen as the principal value. Thus, $\rho = 2\sqrt{2}$, $\theta = \frac{3}{4}\pi$.

ELEMENTARY PROBLEMS OF PLANE ANALYTIC GEOMETRY

§ 5. Projection of a Line Segment. Distance Between Two Points

16. In all our future work we shall always assume that some coordinate system has been chosen. Whenever we speak of any points as given, this will mean that their coordinates are known. If it is required to find some unknown points, we shall consider the problem solved as soon as the coordinates of those points have been calculated.

The present chapter is concerned with the solutions of several elementary problems of analytic geometry.

17. Let there be given an arbitrary segment $\overline{M_1M_2}$ and an axis u (Fig. 11).

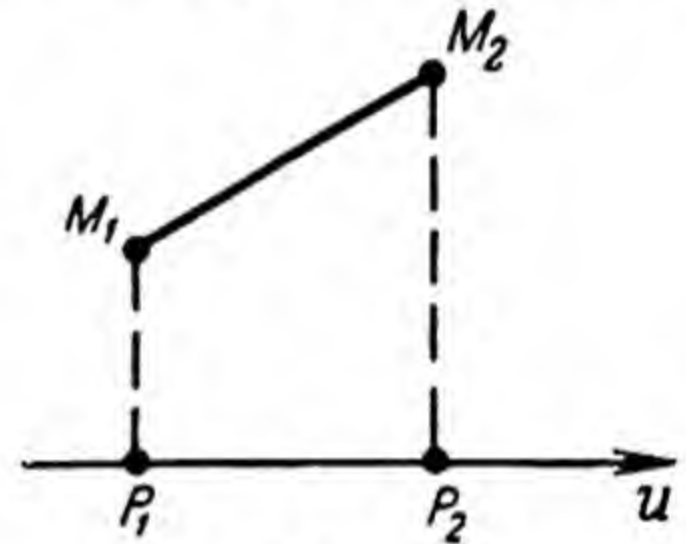


Fig. 11.

From the points M_1 and M_2 , drop perpendiculars to the axis u and denote their respective feet by P_1 and P_2 . Consider the segment $\overline{P_1P_2}$ of the axis u . The *initial* point of $\overline{P_1P_2}$ is the *projection* of the *initial* point of the given segment $\overline{M_1M_2}$; the *terminal* point of $\overline{P_1P_2}$ is the *projection* of the *terminal* point of $\overline{M_1M_2}$. The *value* of the segment $\overline{P_1P_2}$ of the axis u is called the *projection* of the *given segment* $\overline{M_1M_2}$ *on the axis* u ; this is expressed symbolically by the relation

$$\text{proj}_u \overline{M_1M_2} = \overline{P_1P_2}.$$

According to this definition, the *projection of a segment on an axis is a number*, which may be positive (Fig. 11), negative (Fig. 12a), or zero (Fig. 12b).

The necessity of calculating the projections of a segment on the coordinate axes arises particularly frequently in analytic geometry. We shall agree to denote the projection of an arbitrary segment on the axis Ox by the capital letter X , and the projection on the axis Oy by the capital letter Y .

The problem of calculating X , Y when points M_1 , M_2 are given is solved by means of the following

Theorem 3. For any points $M_1 (x_1, y_1)$ and $M_2 (x_2, y_2)$, the projections of the segment $\overline{M_1 M_2}$ on the coordinate axes are expressed by the formulas

$$X = x_2 - x_1, \quad Y = y_2 - y_1. \quad (1)$$

Proof. From the points M_1, M_2 , drop perpendiculars on the axis Ox and denote their feet by P_1, P_2 (Fig. 13). According to

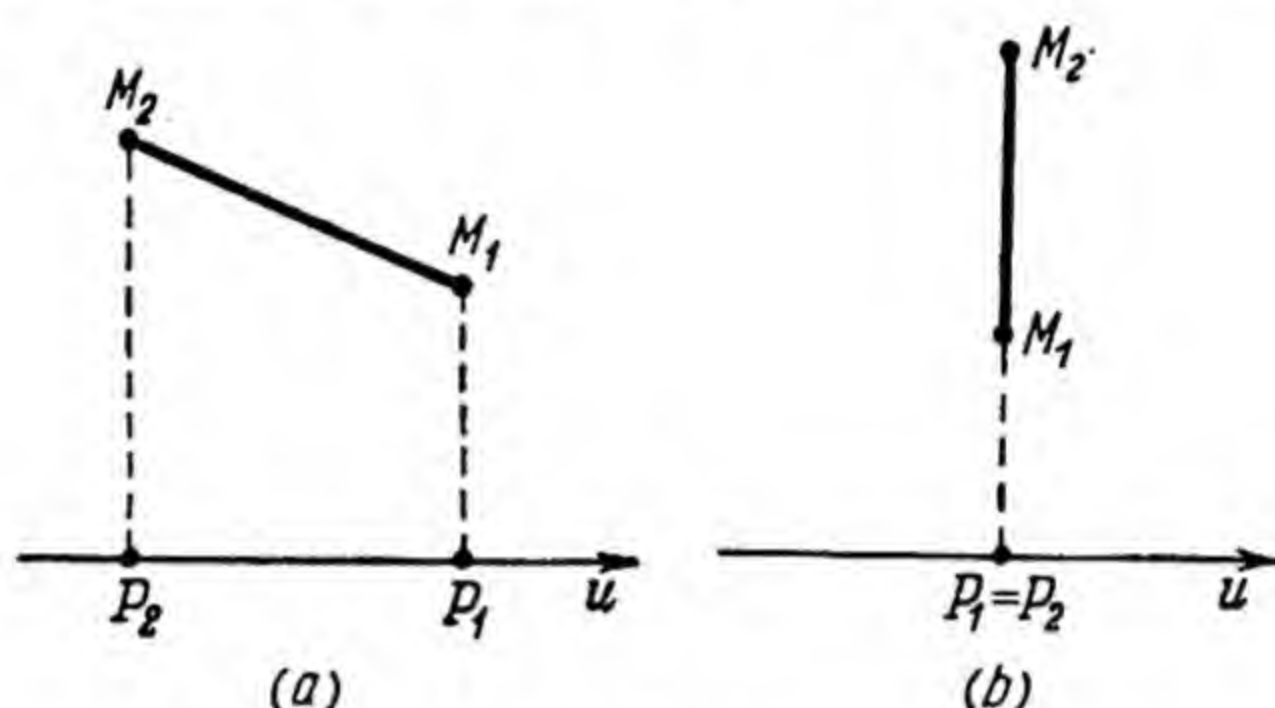


Fig. 12.

Art. 9, the coordinates of these points on the axis Ox are x_1, x_2 , respectively. Hence, by virtue of Theorem 1 of Art. 5,

$$P_1 P_2 = x_2 - x_1.$$

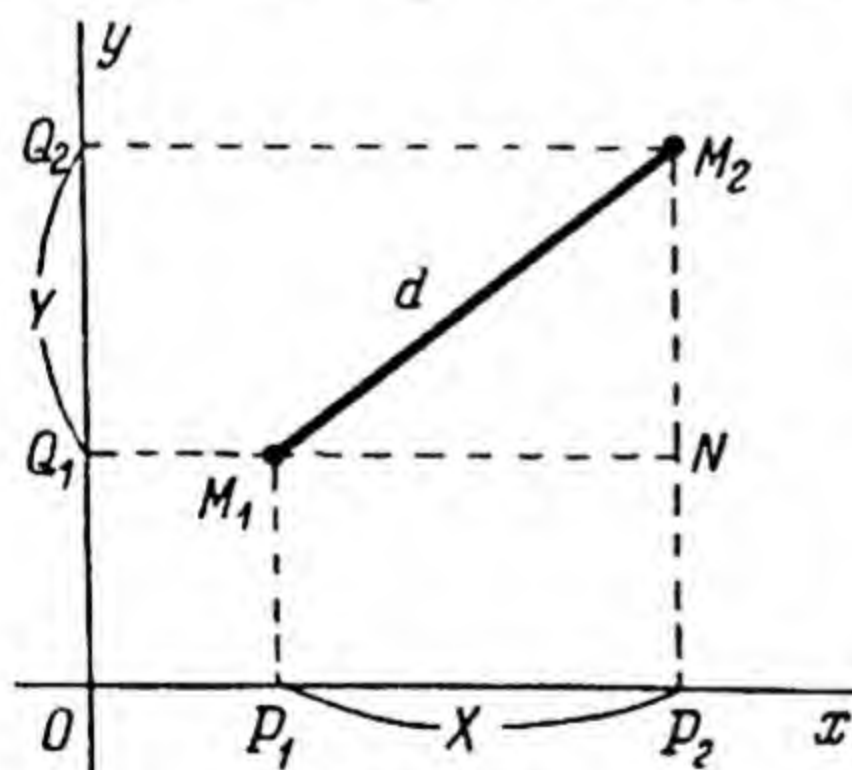


Fig. 13.

But $P_1 P_2 = X$, and therefore $X = x_2 - x_1$. The relation $Y = Q_1 Q_2 = y_2 - y_1$ is established in a similar way. This completes the proof.

Thus, to obtain the projections of a segment on the coordinate axes, subtract the coordinates of its initial point from the corresponding coordinates of its terminal point.

Suppose that the initial point M_1 of the segment coincides with the origin O ; then $x_1 = 0, y_1 = 0$. In this case, we denote the terminal point of the segment simply by the letter M , and the coordinates of the point M by the letters x, y ; by formulas (1), we then obtain

$$X = x, \quad Y = y, \quad (1')$$

where X, Y are the projections of the segment \overline{OM} . The segment \overline{OM} extending from the origin to the given point M is called *the radius vector of that point*. Formulas (1') express the obvious fact that *the rectangular cartesian coordinates of a point are the projections of its radius vector on the coordinate axes*.

18. *The problem of determining the distance between two given points* is one of the most frequently encountered elementary problems of analytic geometry. If the points are given in rectangular cartesian coordinates, the solution of this problem is furnished by the following

Theorem 4. *For any position of points $M_1 (x_1, y_1)$ and $M_2 (x_2, y_2)$ in the plane, the distance d between them is determined by the formula*

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}. \quad (2)$$

Proof. We shall keep the notation used in the preceding theorem, and, in addition, denote by N the intersection of the lines M_1Q_1 and M_2P_2 (Fig. 13). M_1M_2N is a right-angled triangle, and so, by Pythagoras' Theorem,

$$d = \sqrt{M_1N^2 + M_2N^2}.$$

But the lengths of the sides M_1N and M_2N are clearly identical with the absolute values of the projections X, Y of the segment $\overline{M_1M_2}$ on the coordinate axes; therefore

$$d = \sqrt{X^2 + Y^2}.$$

Hence, by Theorem 3, we find

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

as was to be shown.

Example. Find the distance between the points $M_1(-2, 3)$ and $M_2(5, 4)$.

Solution. By formula (2),

$$d = \sqrt{[5 - (-2)]^2 + (4 - 3)^2} = \sqrt{50} = 5\sqrt{2}.$$

19. Let us once more consider the segment $\overline{M_1M_2}$. Through its initial point M , draw a ray u parallel to the axis Ox and having the same direction as Ox (Fig. 14). Denote by θ the angle through which the ray u must be rotated to make its direction coincide with that of the segment $\overline{M_1M_2}$; this angle will be understood as in trigonometry (that is, as a signed quantity determined to within the term $\pm 2n\pi$).

We shall call the angle θ *the polar angle of the segment $\overline{M_1M_2}$ in reference to the given coordinate axes*. Obviously, θ is nothing

more than the polar angle of the point M_2 in the polar coordinate system whose pole is the point M_1 and whose polar axis is the ray u ; in this polar system, the length d of the given segment will be the polar radius of the point M_2 .

Let us now regard the point M_1 as the origin of a new cartesian coordinate system, whose axes are directed similar to the axes of the original cartesian system (in Fig. 14, the new axes are shown in dashed line).

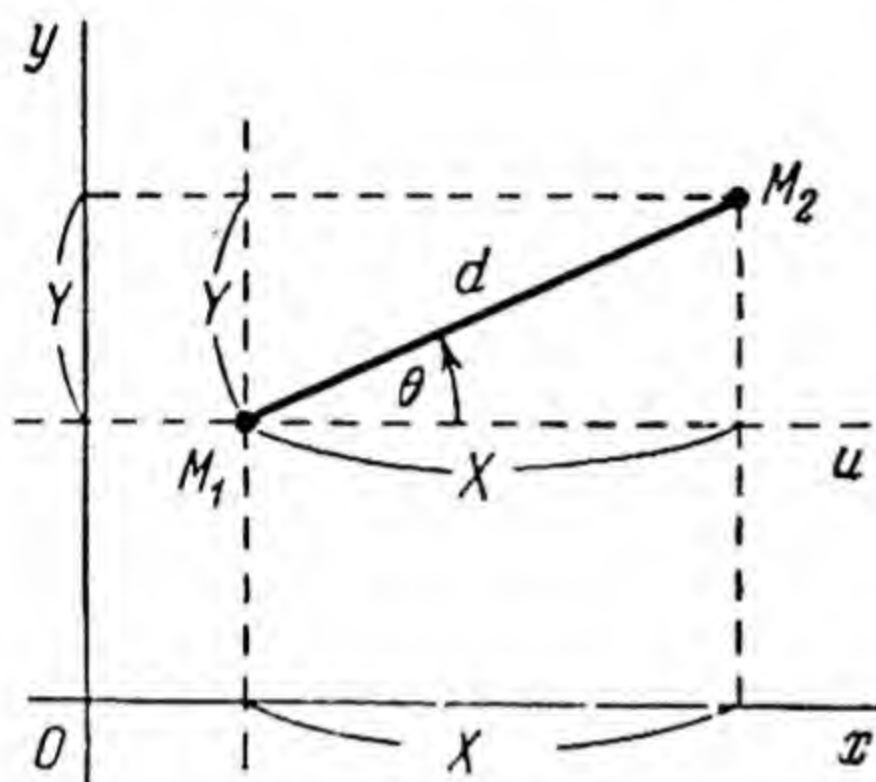


Fig. 14.

The projections of the segment M_1M_2 on the corresponding axes of the old and the new system being identical, we shall denote them, as before, by X , Y . The numbers X , Y are the cartesian coordinates of the point M_2 with respect to the new system. Applying to them formulas (1) of Art. 15, we find

$$X = d \cos \theta, \quad Y = d \sin \theta. \quad (3)$$

Formulas (3) express the projections of an arbitrary segment on the coordinate axes in terms of its length and its polar angle.

From these formulas and in virtue of Theorem 3,

$$x_2 - x_1 = d \cos \theta, \quad y_2 - y_1 = d \sin \theta, \quad (4)$$

or

$$\cos \theta = \frac{x_2 - x_1}{d}, \quad \sin \theta = \frac{y_2 - y_1}{d}. \quad (5)$$

Formulas (5) make it possible to determine the polar angle of a segment from the coordinates of its terminal and initial points (after finding d from formula (2)).

In many cases, the formula

$$\tan \theta = \frac{y_2 - y_1}{x_2 - x_1} \quad (6)$$

can also be conveniently used; this formula is readily derived from (4).

Example 1. Given the length $d = 2\sqrt{2}$ and the polar angle $\theta = 135^\circ$ of a segment, find its projections on the coordinate axes.

Solution. By formulas (3), we find

$$X = 2\sqrt{2} \cos 135^\circ = 2\sqrt{2} \left(-\frac{1}{\sqrt{2}} \right) = -2,$$

$$Y = 2\sqrt{2} \sin 135^\circ = 2\sqrt{2} \frac{1}{\sqrt{2}} = 2.$$

Example 2. Find the polar angle of the segment directed from the point $M_1(5, \sqrt{3})$ to the point $M_2(6, 2\sqrt{3})$.

Solution. By formula (2),

$$d = \sqrt{(6-5)^2 + (2\sqrt{3}-\sqrt{3})^2} = 2.$$

Using formulas (5), we find

$$\cos \theta = \frac{1}{2}, \quad \sin \theta = \frac{\sqrt{3}}{2}.$$

Hence, principal value of θ is 60° .

20. Let u be an *arbitrary* axis, and let φ be the angle of inclination of the segment $\overline{M_1M_2}$ with respect to this axis, that is, the angle through which the axis u has to be rotated to make its direction coincide with that of the segment $\overline{M_1M_2}$.

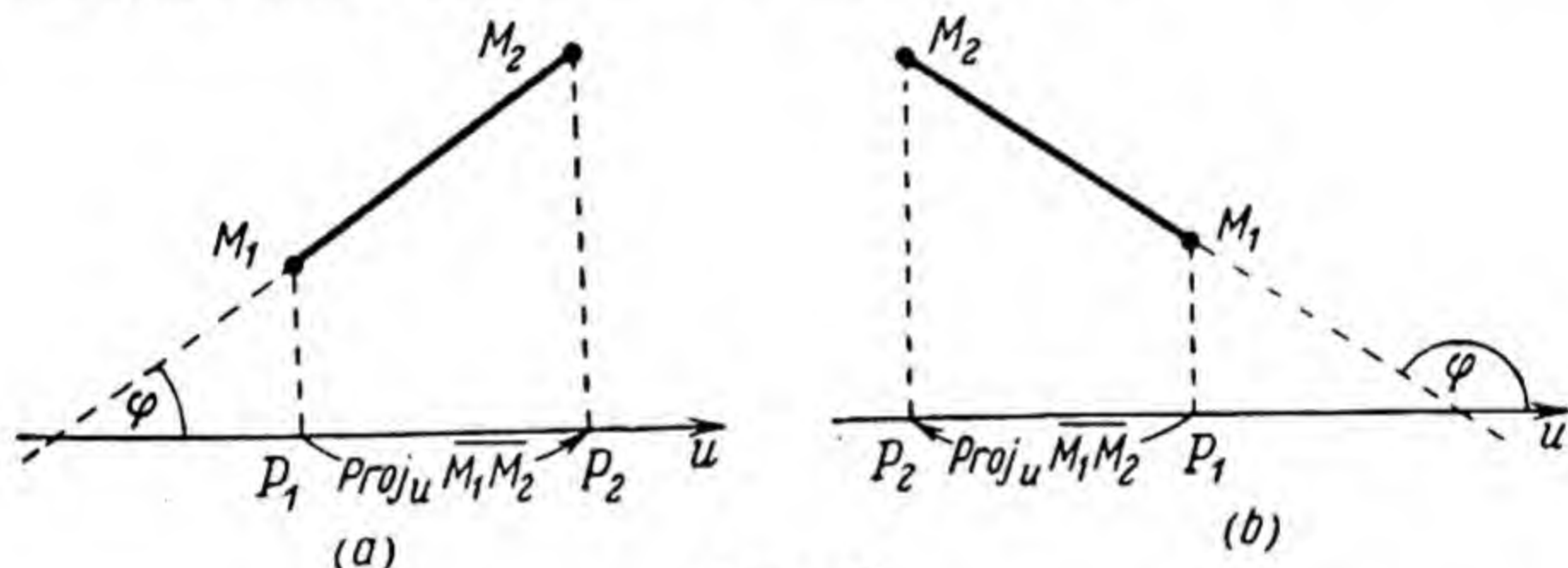


Fig. 15.

The following formula serves for computing the projection of the segment $\overline{M_1M_2}$ on the axis u :

$$\text{proj}_u \overline{M_1M_2} = d \cos \varphi; \quad (7)$$

that is, the projection of a segment on any axis is equal to the length of the segment multiplied by the cosine of its angle of inclination with respect to this axis.

Formula (7) needs no proof, since it is essentially the same as the first of formulas (3) in Art. 19. It should only be noted that, since the sign of an angle does not affect its cosine, the angle φ in formula (7) may be taken as in elementary geometry: in the range 0° to 180° and without regard to sign.

If φ is an acute angle, $\cos \varphi$ and the projection of the segment are positive (Fig. 15a); if φ is an obtuse angle, $\cos \varphi$ and the projection of the segment are negative (Fig. 15b). If φ is a right angle, the projection is zero.

Example. Given the points $M_1(1, 1)$ and $M_2(4, 6)$ Find the projection of the segment $\overline{M_1M_2}$ on the axis passing through the points $A(1, 0)$ and $B(5, 3)$, and directed from A to B .

Solution. Let u be the given axis, φ the angle of inclination of the segment $\overline{M_1M_2}$ with respect to the axis u , θ and θ' the polar angles of the segments $\overline{M_1M_2}$ and \overline{AB} (see Fig. 16, where all these angles have M_1 as vertex).

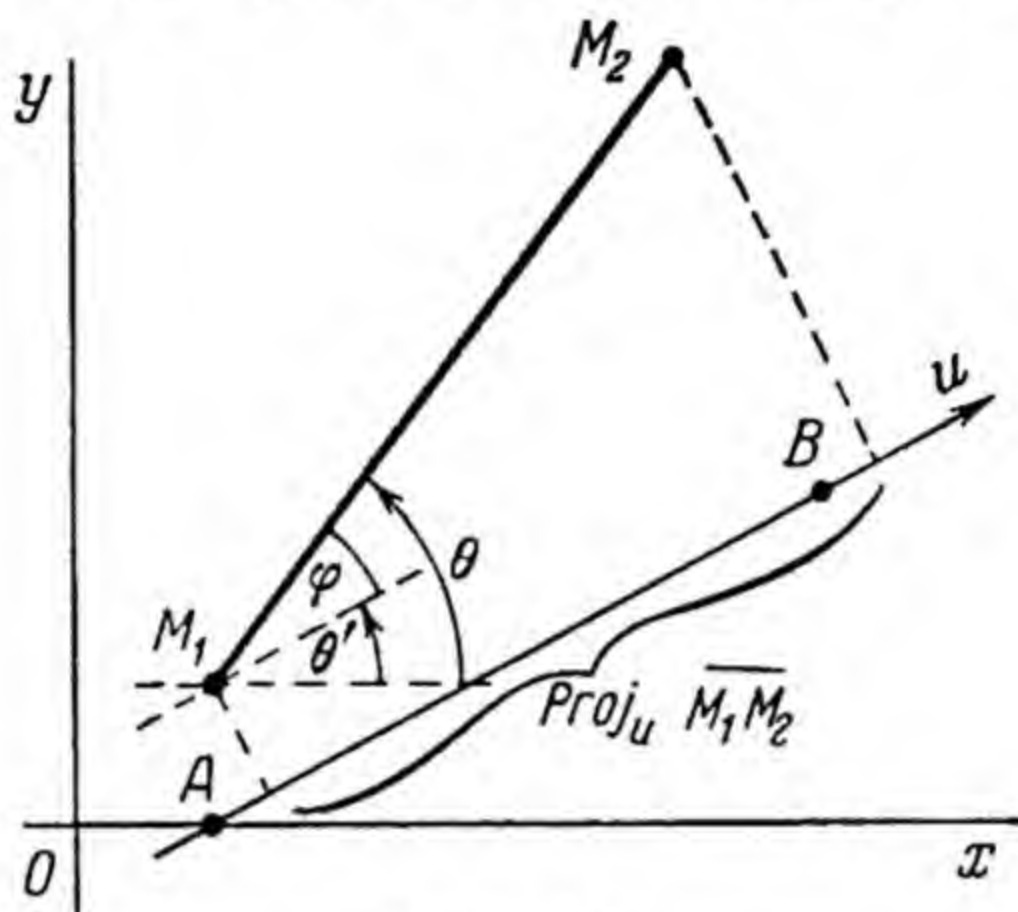


Fig. 16.

It is apparent that $\cos \varphi = \cos (\theta - \theta')$. Denote the projections of the segment $\overline{M_1M_2}$ on the coordinate axes by X, Y , the projections of the segment \overline{AB} by X', Y' , and the lengths of the segments $\overline{M_1M_2}$ and \overline{AB} by d and d' , respectively. By formula (7),

$$\text{proj}_u \overline{M_1M_2} = d \cos \varphi = d \cos (\theta - \theta') = d (\cos \theta \cos \theta' + \sin \theta \sin \theta').$$

Hence, using formulas (3) of Art. 19, we obtain

$$\text{proj}_u \overline{M_1M_2} = d \left(\frac{X}{d} \cdot \frac{X'}{d'} + \frac{Y}{d} \cdot \frac{Y'}{d'} \right) = \frac{XX' + YY'}{d'}.$$

Applying Theorems 3 and 4, we find

$$X = 3, Y = 5, X' = 4, Y' = 3, d' = \sqrt{4^2 + 3^2} = 5.$$

Consequently,

$$\text{proj}_u \overline{M_1M_2} = \frac{3 \cdot 4 + 5 \cdot 3}{5} = \frac{27}{5}.$$

The problem is solved.

§ 6. Calculation of the Area of a Triangle

21. Let there be given three points $A(x_1, y_1)$, $B(x_2, y_2)$ and $C(x_3, y_3)$, which are not on the same straight line. We shall now derive the formula expressing the area S of the triangle ABC in terms of the coordinates of its vertices.

Let φ be the angle between the segments \overline{AB} and \overline{AC} , and let d and d' be the lengths of these segments. As we know from elementary geometry, the area of a triangle is equal to half the product of its two sides and the sine of the included angle; hence

$$S = \frac{1}{2} dd' \sin \varphi. \quad (1)$$

Denote by θ the polar angle of the segment \overline{AB} . If the shortest rotation of the segment \overline{AB} to the segment \overline{AC} through the angle φ is in the positive direction, then, by adding φ to θ , we obtain the polar angle of the segment \overline{AC} ; denoting it by θ' ,

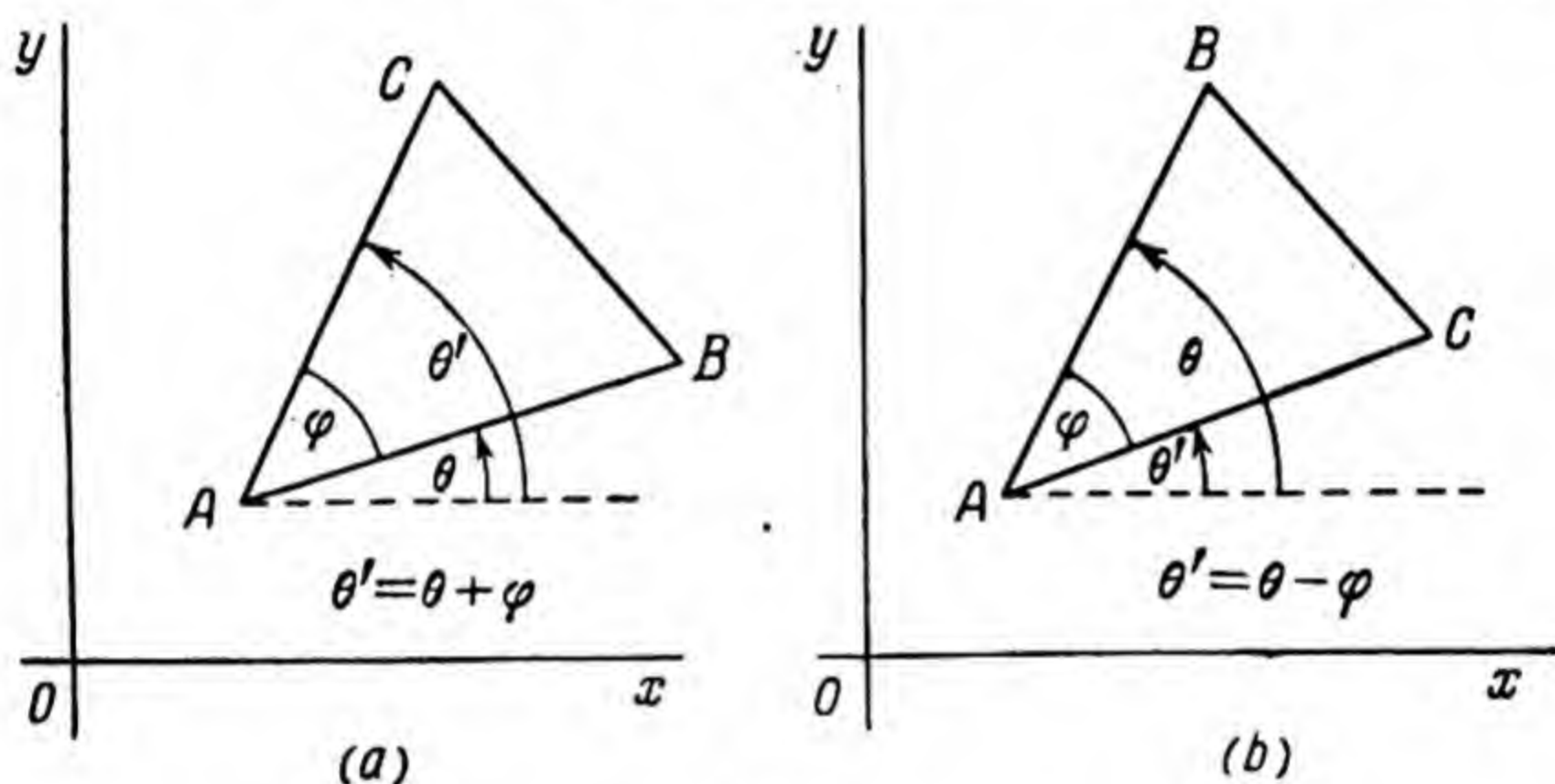


Fig. 17.

we have $\theta' = \theta + \varphi$ (Fig. 17 a). On the other hand, if the shortest rotation of \overline{AB} to \overline{AC} is in the negative direction, then the polar angle θ' of the segment \overline{AC} is obtained by subtracting φ from θ ; in this case, $\theta' = \theta - \varphi$ (Fig. 17 b). Thus, $\varphi = \pm(\theta' - \theta)$; hence, from formula (1), we have

$$S = \pm \frac{1}{2} dd' \sin(\theta' - \theta) = \pm \frac{1}{2} dd' (\sin \theta' \cos \theta - \cos \theta' \sin \theta). \quad (2)$$

Let us denote the projections of the segment \overline{AB} on the coordinate axes by X , Y , and the projections of the segment \overline{AC} by X' , Y' . By formulas (3) of Art. 19,

$$X = d \cos \theta, \quad Y = d \sin \theta.$$

$$X' = d' \cos \theta', \quad Y' = d' \sin \theta'.$$

Removing the parentheses in the right-hand member of (2) and using these last relations, we find

$$S = \pm \frac{1}{2} (XY' - X'Y). \quad (3)$$

According to Theorem 3 of Art. 17,

$$\begin{aligned} X &= x_2 - x_1, & Y &= y_2 - y_1, \\ X' &= x_3 - x_1, & Y' &= y_3 - y_1. \end{aligned}$$

Substituting these expressions in formula (3), we obtain

$$S = \pm \frac{1}{2} [(x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1)]. \quad (4)$$

Since the expression within the brackets is a determinant of the second order ^{*}), formula (4) may also be written as

$$\pm S = \frac{1}{2} \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix}. \quad (5)$$

This result may be expressed in the form of the following

Theorem 5. *For any three points $A(x_1, y_1)$, $B(x_2, y_2)$, $C(x_3, y_3)$, not lying on the same line, the area S of the triangle ABC is determined by formula (5). The right-hand member of this formula is equal to $+S$ if the shortest rotation of the segment \overline{AB} to the segment \overline{AC} is in the positive direction; it is equal to $-S$ if the shortest rotation of \overline{AB} to \overline{AC} is in the negative direction.*

Example. Given the points $A(1, 1)$, $B(6, 4)$, $C(8, 2)$. Find the area of the triangle ABC .

Solution. By formula (5),

$$\frac{1}{2} \begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} = \frac{1}{2} \begin{vmatrix} 5 & 3 \\ 7 & 1 \end{vmatrix} = -8 = -S.$$

Hence, $S = 8$. The fact that the right-hand member of (5) is negative in the present case means that the shortest rotation of \overline{AB} to \overline{AC} is in the negative direction.

§ 7. Division of a Line Segment in a Given Ratio

22. The problem of dividing a line segment in a given ratio belongs to those elementary problems of analytic geometry which find numerous applications. Before giving a precise formulation

^{*} See the Appendix (page 225) for basic information on determinants.

of the problem, we must explain at some length what is meant by *the ratio in which a point divides a given segment*.

Let there be given any two distinct points in the plane, and let one of these points be designated as the first, and the other point as the second. Denote them, in this order, by M_1 and M_2 . Through these points, draw a straight line u and assign to it a positive direction, thereby making it an axis.

Further, let M be a third point of the axis u , occupying any position on the axis, provided only that it does not coincide with the point M_2 (Fig. 18).

The number λ determined by the relation

$$\lambda = \frac{M_1M}{MM_2}, \quad (1)$$

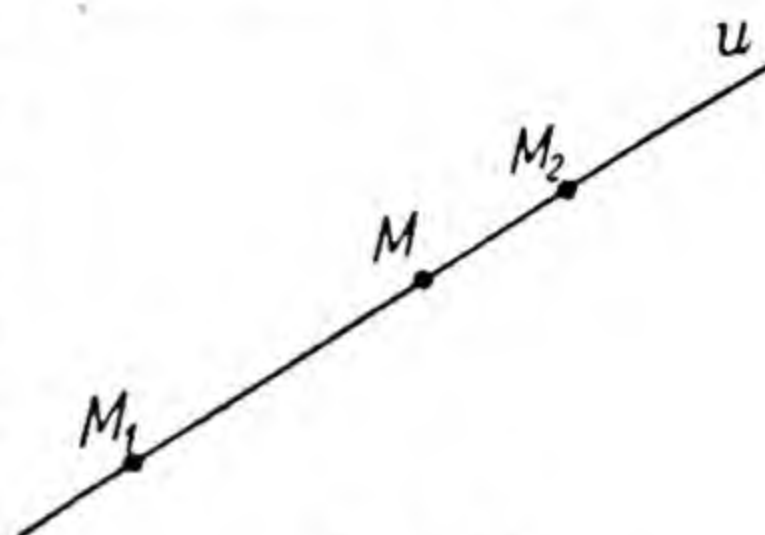


Fig. 18.

where $\overline{M_1M}$ and $\overline{MM_2}$ are the values of the directed segments $\overline{M_1M}$ and $\overline{MM_2}$ of the axis u , is called the ratio in which the point M divides the directed segment $\overline{M_1M_2}$.

Note 1. The number λ does not depend upon the choice of the positive direction on the line u determined by the points M_1 and M_2 . For, if the positive direction of this line is reversed, then the values $\overline{M_1M}$ and $\overline{MM_2}$ will simultaneously change their signs (without changing their moduli), so that the fraction $\frac{\overline{M_1M}}{\overline{MM_2}}$ will obviously remain unchanged.

Note 2. Moreover, the number λ does not depend upon the choice of the scale for measurement of lengths. For, if the scale is changed, the values of all segments on the axis M_1M_2 will be multiplied by the same factor, and so the ratio $\frac{\overline{M_1M}}{\overline{MM_2}}$ will remain unchanged.

Note 3. If the condition that M must not coincide with M_2 is not imposed, then in the case when M coincides with M_2 , relation (1) does not determine any number (because $\overline{MM_2} = 0$). The ratio $\frac{\overline{M_1M}}{\overline{MM_2}}$ is then said (for a reason explained in the next article) to be "infinite".

23. Suppose that the positive direction on the line M_1M_2 is chosen so that the segment $\overline{M_1M_2}$ is positively directed; then $\overline{M_1M_2}$ is a positive number. If, then, the point M lies *between the points*

M_1 and M_2 , the numbers M_1M , MM_2 are also positive, so that the ratio $\lambda = \frac{M_1M}{MM_2}$ is a *positive* number. As the point M approaches the point M_1 , λ approaches zero (λ becomes zero when M coincides with M_1); as the point M approaches the point M_2 , λ increases without limit and is said to tend to infinity (we therefore say that λ "becomes infinite" when M coincides with M_2).

Suppose now that the point M (on the line determined by the points M_1 and M_2) lies *outside the segment* $\overline{M_1M_2}$. In this case, one of the numbers M_1M , MM_2 is positive, the other negative, and $\lambda = \frac{M_1M}{MM_2}$ is a *negative* number since M_1M and MM_2 differ in sign.

24. In analytic geometry, the problem of dividing a line segment in a given ratio is as follows:

Given the coordinates of two points $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ and the ratio λ in which some (unknown) point M divides the segment $\overline{M_1M_2}$, to find the coordinates of M .

The solution of this problem is furnished by the following

Theorem 6. If a point $M(x, y)$ divides the segment $\overline{M_1M_2}$ in the ratio λ , then the coordinates of that point are given by the formulas

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}, \quad y = \frac{y_1 + \lambda y_2}{1 + \lambda}. \quad (2)$$

Proof. Project the points M_1 , M_2 and M on the axis Ox and denote their projections by P_1 , P_2 and P , respectively (Fig. 19). By a well-known theorem of elementary geometry, which states that parallel lines intercept proportional segments on transversals, we have

$$\frac{P_1P}{PP_2} = \frac{M_1M}{MM_2} = \lambda. \quad (3)$$

According to Theorem 3 (Art. 17),

$$P_1P = x - x_1, \quad PP_2 = x_2 - x$$

Hence, from (3), we get

$$\frac{x - x_1}{x_2 - x} = \lambda.$$

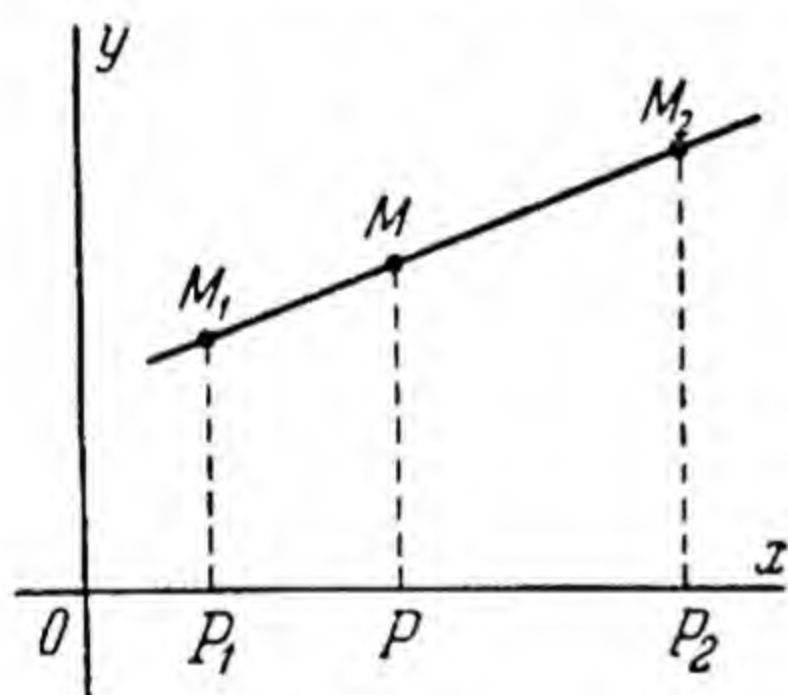


Fig. 19.

Solving this equation for the unknown x , we find that

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}.$$

Thus, we have obtained the first of formulas (2). The second formula is established in an entirely analogous manner, by projecting the same points on the axis Oy ; this completes the proof.

Note. If $\lambda = -1$, formulas (2) become meaningless, since the denominators of these formulas are then zero ($1 + \lambda = 0$). But in this case the problem itself has no solution; in fact, no point can divide the segment $\overline{M_1M_2}$ in the ratio $\lambda = -1$, for, if $\frac{M_1M}{MM_2} = -1$, then $M_1M = -MM_2$, or $M_1M + MM_2 = M_1M_2 = 0$, which is impossible, since M_1, M_2 are, by hypothesis, distinct points.

25. Note an important special case of the theorem just proved: If $M_1 (x_1, y_1)$ and $M_2 (x_2, y_2)$ are two arbitrary points and $M (x, y)$ is the midpoint of the segment $\overline{M_1M_2}$, then

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}.$$

These formulas are obtained from formulas (2) of Art. 24 by setting $\lambda = 1$ *). Accordingly, we may state that *each coordinate of the midpoint of a segment is equal to half the sum of the corresponding coordinates of its endpoints.*

Example 1. Given the points $M_1 (1, 1)$ and $M_2 (7, 4)$. On the line through M_1 and M_2 , find the point M which is two times closer to M_1 than to M_2 and lies between the points M_1 and M_2 .

Solution. The required trisection point divides the segment $\overline{M_1M_2}$ in the ratio $\lambda = \frac{1}{2}$. By formulas (2) of Art. 24, the coordinates of M are: $x = 3, y = 2$.

Example 2. Given the points $M_1 (1, 1)$ and $M_2 (7, 4)$. On the line through M_1 and M_2 , find the point M which is two times closer to M_1 than to M_2 and lies outside the segment bounded by the points M_1 and M_2 .

Solution. The required point divides the segment $\overline{M_1M_2}$ in the ratio $\lambda = -\frac{1}{2}$.

By formulas (2) Art. 24, the coordinates of M are: $x = -5, y = -2$.

Example 3. Given the vertices $A(5, -1), B(-1, 7), C(1, 2)$ of a triangle. Find the length of the internal bisector of the angle A in this triangle.

Solution. Let M be the point of intersection of the bisector with side BC , and let c and b be the respective lengths of the sides AB and AC . From elementary geometry, the bisector of an angle of a triangle divides the opposite side in

*) If M is the midpoint of the segment $\overline{M_1M_2}$, then $M_1M = MM_2$, and hence $\lambda = \frac{M_1M}{MM_2} = 1$.

the ratio of the sides containing the angle. Thus, the point M divides the segment \overline{BC} in the ratio

$$\lambda = \frac{BM}{MC} = \frac{c}{b}.$$

Using formula (2) of Art. 18, we find the lengths of the sides AB and AC :

$$c = \sqrt{(5+1)^2 + (-1-7)^2} = 10, \quad b = \sqrt{(5-1)^2 + (-1-2)^2} = 5.$$

Hence, $\lambda=2$. Applying formulas (2) of Art. 24, we find the coordinates of the point M : $x = \frac{1}{3}$, $y = \frac{11}{3}$.

Using formula (2) of Art. 18 once again, we obtain the required length of the bisector $AM = \frac{14}{3}\sqrt{2}$.

Example 4. The masses m_1, m_2, m_3 are placed at the points $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$, respectively. Find the centre of gravity of this system of masses.

Solution. Let us first find the centre of gravity $M'(x', y')$ for the system of two masses m_1 and m_2 . According to a well-known mechanical principle, the centre of gravity of this system of masses divides the segment $\overline{M_1M_2}$ into parts inversely proportional to the masses m_1, m_2 , that is, in the ratio $\lambda = \frac{m_2}{m_1}$. By formulas (2) of Art. 24, we find

$$x' = \frac{x_1 + \frac{m_2}{m_1} x_2}{1 + \frac{m_2}{m_1}} = \frac{x_1 m_1 + x_2 m_2}{m_1 + m_2},$$

$$y' = \frac{y_1 + \frac{m_2}{m_1} y_2}{1 + \frac{m_2}{m_1}} = \frac{y_1 m_1 + y_2 m_2}{m_1 + m_2}.$$

Let $M(x, y)$ be the centre of gravity of the system of three masses m_1, m_2, m_3 . The position of the point M will remain unchanged if the masses m_1, m_2 are concentrated at the point M' . In other words, the point M is the centre of gravity of the system of the following two masses: the mass m_3 placed at the point M_3 , and the mass $m_1 + m_2$ placed at the point M' . We can therefore find the required point M as the point dividing the segment $\overline{M'M_3}$ in the ratio $\lambda = \frac{m_3}{m_1 + m_2}$. Applying formulas (2) of Art. 24, we obtain:

$$x = \frac{x' + \frac{m_3}{m_1 + m_2} x_3}{1 + \frac{m_3}{m_1 + m_2}} = \frac{\frac{x_1 m_1 + x_2 m_2}{m_1 + m_2} + \frac{m_3}{m_1 + m_2} x_3}{1 + \frac{m_3}{m_1 + m_2}} = \frac{x_1 m_1 + x_2 m_2 + x_3 m_3}{m_1 + m_2 + m_3},$$

$$y = \frac{y' + \frac{m_3}{m_1 + m_2} y_3}{1 + \frac{m_3}{m_1 + m_2}} = \frac{\frac{y_1 m_1 + y_2 m_2}{m_1 + m_2} + \frac{m_3}{m_1 + m_2} y_3}{1 + \frac{m_3}{m_1 + m_2}} = \frac{y_1 m_1 + y_2 m_2 + y_3 m_3}{m_1 + m_2 + m_3}.$$

Note. Given the system of masses m_1, m_2, \dots, m_k placed at the points $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_k(x_k, y_k)$, the coordinates of the centre of gravity of this system of masses are determined by the formulas

$$x = \frac{x_1 m_1 + x_2 m_2 + \dots + x_k m_k}{m_1 + m_2 + \dots + m_k},$$

$$y = \frac{y_1 m_1 + y_2 m_2 + \dots + y_k m_k}{m_1 + m_2 + \dots + m_k}.$$

To prove this, the reader should use formulas (2) of Art. 24 and the method of mathematical induction.

§ 8. Transformation of Cartesian Coordinates by Translation of Axes

26. In problems of analytic geometry, as we already know, the position of the given geometric objects is determined in reference to some coordinate system. It may, however, be necessary to replace the given reference system by another coordinate system, thought to be more convenient for some reason. But an arbitrary point has, in general, different coordinates in different coordinate systems. Therefore, when making use of two coordinate systems in a single problem, the following necessity arises: given the coordinates of an arbitrary point in one system, to calculate the coordinates of that point in reference to the other system. This is achieved by means of the *coordinate transformation formulas* corresponding to the required change of coordinate system.

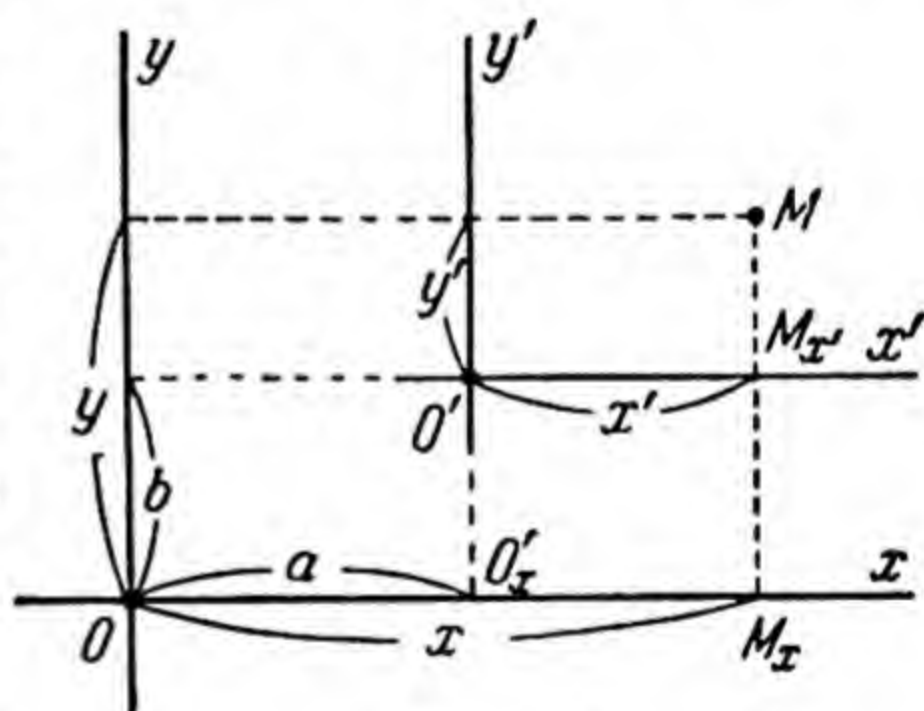


Fig. 20.

27. Let us, first of all, establish the *formulas for transformation of cartesian coordinates by translation of axes*, that is, the formulas corresponding to the change of a cartesian coordinate system when the origin is moved to a new position without changing the direction of the axes (and the scale).

Let Ox, Oy be the old coordinate axes, and let $O'x', O'y'$ be the new axes (Fig. 20). The position of the new axes relative to the old system is determined by giving the old coordinates of the new origin: $O' (a, b)$. We shall refer to the number a as the amount of the shift in the direction of the axis Ox and to the number b as the amount of the shift in the direction of the axis

Oy. An arbitrary point M of the plane has coordinates (x, y) with respect to the old axes; the same point M has different coordinates (x', y') with respect to the new axes. We now propose to establish the formulas expressing x, y in terms of x', y' (or x', y' in terms of x, y).

Project the point O' on the axis Ox , and the point M on the axes Ox and $O'x'$.

Denote the projection of the point O' on the axis Ox by O'_x and the projections of the point M on the axes Ox and $O'x'$ by M_x and $M_{x'}$. Obviously, the value of the segment $O'_x M_x$ of the axis Ox is equal to the value of the segment $O' M_{x'}$ of the axis $O'x'$. But $O' M_{x'} = x'$; consequently, $O'_x M_x = x'$. Also, $OO'_x = a$, $OM_x = x$. By the fundamental identity (see Art. 3), $OM_x = OO'_x + O'_x M_x$; hence, from the foregoing, we have $x = x' + a$. Similarly, by projecting O' and M on the axes Oy and $O'y'$, we find $y = y' + b$.

Thus,

$$x = x' + a, \quad y = y' + b. \quad (1)$$

These are the required formulas, which may also be written in the form

$$x' = x - a, \quad y' = y - b. \quad (1')$$

This result can be formulated as follows: *When a cartesian coordinate system is translated by an amount a in the direction of the axis Ox , and by an amount b in the direction of the axis Oy , this means the subtraction of a from the abscissas, and of b from the ordinates of all points.*

§ 9. Transformation of Rectangular Cartesian Coordinates by Rotation of Axes

28. We now proceed to establish the *formulas for transformation of rectangular cartesian coordinates under rotation of axes*, that is, the formulas corresponding to the change of a rectangular cartesian system when both axes are turned in the same direction and through the same angle, without changing the origin and the scale.

Let Ox, Oy be the old, and Ox', Oy' the new coordinate axes (Fig. 21). The position of the new axes relative to the old system is determined by giving the angle of rotation which brings the old axes into coincidence with the new axes. This angle will be denoted by the letter α and understood as in trigonometry; the positive direction of rotation will be defined as in Art. 15.

An arbitrary point M of the plane has coordinates (x, y) with respect to the old axes; the same point has, in general, different coordinates (x', y') with respect to the new axes. Namely, $x = OM_x$, $y = OM_y$, $x' = OM_{x'}$, $y' = OM_{y'}$ (see Fig. 21). Our objective is to establish the formulas expressing x', y' in terms of x, y (or x, y in terms of x', y').

Let (ρ, θ) be the polar coordinates of the point M , when taking Ox as the polar axis, and let (ρ, θ') be the polar coordinates of the same point M , when taking Ox' as the polar axis (in either case, $\rho = |OM|$). Obviously, $\theta = \theta' + \alpha$. Further, by formulas (1) of Art. 15,

$$x = \rho \cos \theta, \quad y = \rho \sin \theta,$$

and similarly,

$$x' = \rho \cos \theta', \quad y' = \rho \sin \theta'.$$

Hence,

$$\begin{aligned} x &= \rho \cos \theta = \rho \cos (\theta' + \alpha) = \rho (\cos \theta' \cos \alpha - \sin \theta' \sin \alpha) = \\ &= \rho \cos \theta' \cos \alpha - \rho \sin \theta' \sin \alpha = x' \cos \alpha - y' \sin \alpha, \\ y &= \rho \sin \theta = \rho \sin (\theta' + \alpha) = \rho (\cos \theta' \sin \alpha + \sin \theta' \cos \alpha) = \\ &= \rho \cos \theta' \sin \alpha + \rho \sin \theta' \cos \alpha = x' \sin \alpha + y' \cos \alpha. \end{aligned}$$

Thus

$$\left. \begin{aligned} x &= x' \cos \alpha - y' \sin \alpha, \\ y &= x' \sin \alpha + y' \cos \alpha. \end{aligned} \right\} \quad (1)$$

These are the required formulas, that is, the formulas expressing, for a rotation of axes through an angle α , the old coordinates (x, y) of an arbitrary point M in terms of the new coordinates (x', y') of M .

The formulas expressing the new coordinates x', y' of the point M in terms of its old coordinates x, y can be derived from relations (1), regarded as a system of two equations in the two unknowns x', y' , by solving this system for x', y' . However, these formulas can also be obtained immediately by the following argument: *If the new system is obtained by a rotation of the old system through an angle α , then the old system is obtained by a rotation of the new system through the angle $-\alpha$; we may therefore interchange the old and the new coordinates in relations (1),*

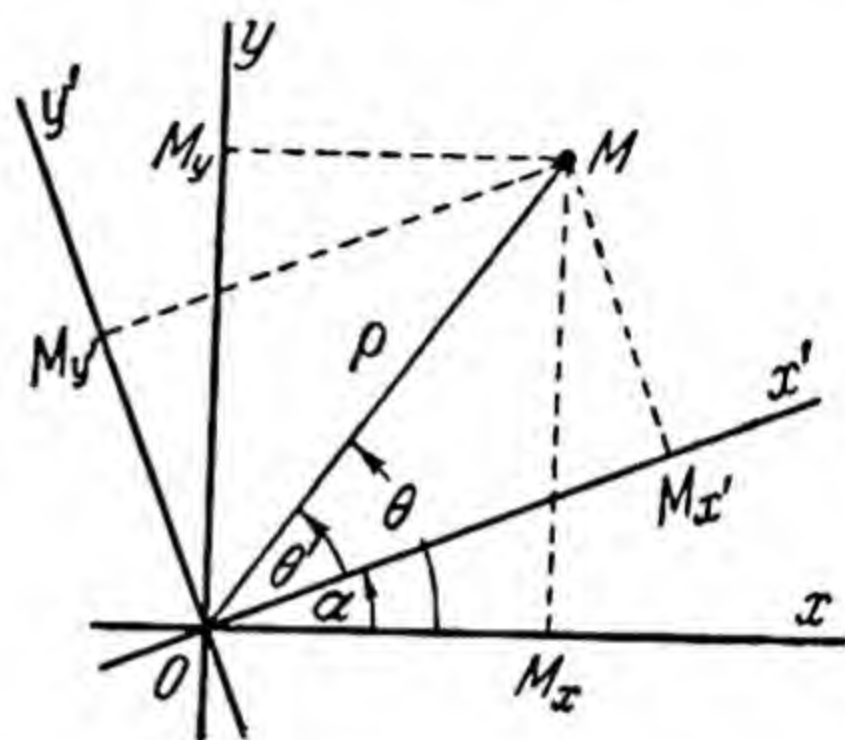


Fig. 21.

simultaneously replacing α by $-\alpha$. On performing this transformation, we get

$$\left. \begin{aligned} x' &= x \cos \alpha + y \sin \alpha, \\ y' &= -x \sin \alpha + y \cos \alpha, \end{aligned} \right\} \quad (2)$$

which are the required formulas.

§ 10. Transformation of Rectangular Cartesian Coordinates by Change of Origin and Rotation of Axes

29. We shall now consider a motion of axes, which may be achieved by a *translation followed by a rotation*. (The scale is assumed to remain unchanged.)

Let a denote the shift of the system in the direction of the axis Ox , and b the shift in the direction of the axis Oy ; let α be the angle of rotation of the system. Denote the new axes by $O'x'$ and $O'y'$. An arbitrary point M of the plane has coordinates (x, y) with respect to the old axes; the same point M has, in general,

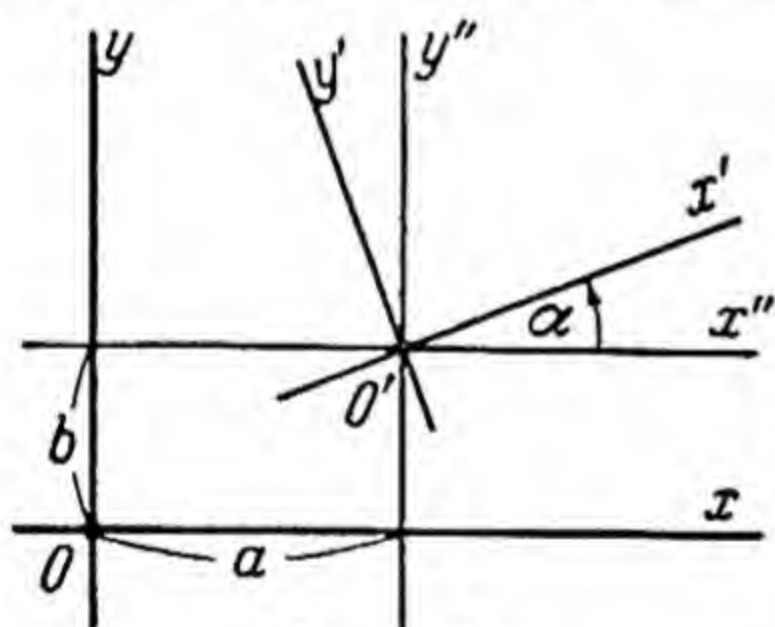


Fig. 22.

different coordinates (x', y') with respect to the new axes. We propose to find the formulas expressing x', y' in terms of x, y , as well as the formulas expressing x, y in terms of x', y' .

To achieve this, we introduce an auxiliary coordinate system, whose axes have the same directions as the axes of the old system, and whose origin coincides with that of the new system (Fig. 22); let $O'x''$, $O'y''$ be the axes of the auxiliary system, and let x'', y'' be the coordinates of the point M with respect to these axes. Our auxiliary system is obtained by translating the old system a units in the direction of Ox and b units in the direction of Oy , so that, by Art. 27,

$$\begin{aligned} x &= x'' + a, \\ y &= y'' + b. \end{aligned}$$

Further, the new system is obtained by rotating the auxiliary system through an angle α ; therefore, by Art. 28,

$$\begin{aligned} x'' &= x' \cos \alpha - y' \sin \alpha, \\ y'' &= x' \sin \alpha + y' \cos \alpha. \end{aligned}$$

Substituting these expressions for x'' , y'' in the right-hand members of the preceding relations, we have

$$\left. \begin{aligned} x &= x' \cos \alpha - y' \sin \alpha + a, \\ y &= x' \sin \alpha + y' \cos \alpha + b. \end{aligned} \right\} \quad (1)$$

Solving the system (1) for x' and y' , we find

$$\left. \begin{aligned} x' &= (x - a) \cos \alpha + (y - b) \sin \alpha, \\ y' &= -(x - a) \sin \alpha + (y - b) \cos \alpha. \end{aligned} \right\} \quad (2)$$

The last two pairs of relations are the formulas we have been seeking.

Formulas (1) express the *old* coordinates of an arbitrary point *in terms* of its *new* coordinates; on the other hand, formulas (2) express the *new* coordinates *in terms* of the *old* coordinates.

We shall formulate this result as the following

Theorem 7. *If the axes of a rectangular cartesian system are translated a units in the direction of Ox and b units in the direction of Oy , and if, in addition, the axes are rotated through an angle α (the scale remaining unchanged), then the resulting change of system is represented by the coordinate transformation formulas*

$$\left. \begin{aligned} x &= x' \cos \alpha - y' \sin \alpha + a, \\ y &= x' \sin \alpha + y' \cos \alpha + b, \end{aligned} \right\} \quad (1)$$

which express the old coordinates x , y of an arbitrary point in the plane in terms of its new coordinates x' , y' , and by the formulas

$$\left. \begin{aligned} x' &= (x - a) \cos \alpha + (y - b) \sin \alpha, \\ y' &= -(x - a) \sin \alpha + (y - b) \cos \alpha, \end{aligned} \right\} \quad (2)$$

which are algebraically equivalent to (1) and express the new coordinates in terms of the old.

Example. Find the coordinate transformation formulas corresponding to a shift of the origin to the point O' (2, 3) and a rotation of the axes through $+45^\circ$.

Solution. Letting $a = 2$, $b = 3$, $\alpha = \frac{\pi}{4}$ in formulas (1), we get the following expressions for the old coordinates in terms of the new:

$$\begin{aligned} x &= \frac{x' - y'}{\sqrt{2}} + 2, \\ y &= \frac{x' + y'}{\sqrt{2}} + 3. \end{aligned}$$

Hence (or from formulas (2)), we obtain the expressions for the new coordinates in terms of the old:

$$\begin{aligned}x' &= \frac{x+y}{\sqrt{2}} - \frac{5}{\sqrt{2}}, \\y' &= \frac{-x+y}{\sqrt{2}} - \frac{1}{\sqrt{2}}.\end{aligned}$$

Note. The usual position of the coordinate axes in diagrams is such that the shortest rotation of the positive x -axis into the positive y -axis will be in the *counterclockwise* direction. In this case, the coordinate system is called *right-handed*. Sometimes, however, use is made of a system whose axes are positioned in a different manner, namely, so that the shortest rotation of the positive x -axis into the positive y -axis will be in the *clockwise* direction, in which case the coordinate system is called *left-handed*.

Let there be given two (rectangular cartesian) coordinate systems. If they are *both right-handed*, or *both left-handed*, then the axes of one system can be brought into coincidence with the axes of the other by means of a translation followed by a rotation through a certain angle; hence and from the foregoing, it follows that, when replacing one of these systems by the other, the coordinates of any point in the plane are transformed according to formulas of the form (1). If, on the other hand, *one of the given systems is right-handed and the other left-handed*, then the axes of one system cannot be brought into coincidence with those of the other by a translation and a subsequent rotation; in fact, if the positive semi-axis of abscissas of one (the "old") system is carried, by a translation and a rotation, into the positive semi-axis of abscissas of the other (the "new") system, then their positive semi-axes of ordinates will go in opposite directions. Consequently, when replacing one of these systems by the other, the coordinates are transformed according to the formulas obtained from (1) by changing the sign of y' . Thus, the *general* formulas for transformation of rectangular cartesian coordinates (provided that the scale remains unchanged) may be written as

$$\left. \begin{aligned}x &= x' \cos \alpha \mp y' \sin \alpha + a, \\y &= x' \sin \alpha \pm y' \cos \alpha + b,\end{aligned} \right\} \quad (3)$$

where a , b are the old coordinates of the new origin, and α is the angle through which the old axis of abscissas must be rotated to go into the new axis of abscissas. In formulas (3), the *upper* signs correspond to the case when the change is made from a *right-handed system to another right-handed one*, or from a *left-handed system to another left-handed one*; the *lower* signs correspond to the case when *one of the systems is right-handed, and the other left-handed*. Also, it must be borne in mind that, if the old system is a left-handed one, the angle α is measured positively in the clockwise direction.

THE EQUATION OF A CURVE

§ 11. The Concept of the Equation of a Curve. Examples of Curves Represented by Equations

30. In elementary geometry, only a small number of curves (the straight line, circle, broken lines) are subjected to a detailed investigation. The needs of engineering, however, pose before mathematics the general problem of investigating various curves of diverse shape and properties. To solve this general problem, more advanced methods are required than those of elementary geometry. Such advanced methods are furnished by algebra and mathematical analysis. The use of the methods of algebra and analysis is based on a uniform mode of determining a curve, namely, that of representing a curve by an *equation*.

31. Let x and y be two arbitrary variables. This means that both the symbol x and the symbol y represent any (real) numbers whatsoever. A relation of the form $F(x, y) = 0$, where $F(x, y)$ denotes an expression containing x and y , is called an *equation* in two variables, x and y (provided that $F(x, y) = 0$ is valid not identically, that is, *not for every pair of numbers* x, y). Examples of equations are $2x + 7y - 1 = 0$, $x^2 + y^2 - 25 = 0$, $\sin x + \sin y - 1 = 0$, etc.

If the relation $F(x, y) = 0$ is valid for *all* values of x, y , it is called an *identity*. Examples of identities are $(x + y)^2 - x^2 - 2xy - y^2 = 0$, $(x + y)(x - y) - x^2 + y^2 = 0$, etc.

The left-hand members of equations in two variables, occurring in the succeeding pages, may contain other symbols a, b, c, \dots , apart from x and y ; but in such cases we shall assume these other symbols to be fixed (though perhaps unspecified) numbers, and we shall call them the constant *parameters* of a given equation. For example, the equation $ax + by - 1 = 0$ has a and b as its parameters.

32. Two numbers $x = x_0, y = y_0$ are said to *satisfy* an equation in two variables if the equation holds true when these numbers are substituted in it for the variables. For instance, the numbers $x = 3, y = 4$ satisfy the equation $x^2 + y^2 - 25 = 0$, because its left-hand member vanishes upon substitution of these numbers;

on the other hand, the numbers $x = 1$, $y = 2$ do not satisfy this equation, since their substitution in the left-hand member does not make it zero.

33. Consider an arbitrary equation $F(x, y) = 0$. Let x and y denote numbers *satisfying* this equation (rather than any arbitrary numbers). In general, x and y may still vary under this condition; but they may no longer vary in an arbitrary manner with respect to each other, because *the possible values of y are determined by assigning a value to x* . The equation $F(x, y) = 0$ is therefore said to establish a *functional relation* between the variables x and y .

34. The fundamental concept of analytic geometry is that of *the equation of a curve*. We shall now explain the meaning of this concept.

Let there be given any curve in the plane; also, let a coordinate system be chosen.

The equation of a given curve (in a chosen coordinate system) is defined as the equation $F(x, y) = 0$ in two variables which is satisfied by the coordinates x, y of all points lying on the curve and by the coordinates of no other point.

Thus, if the *equation* of a curve is known, we can determine for each point of the plane whether it lies on that curve or not. To answer this question, it is necessary merely to substitute the coordinates of the point for the variables in the equation. If the coordinates of the point under test *satisfy* the equation, then the point *lies* on the curve; if they *do not satisfy* the equation, then the point *does not lie* on the curve.

The definition just made constitutes the basis of the methods of analytic geometry, which consist essentially in *the investigation of curves by analysing their equations*.

35. In many problems the equation of a curve is regarded as something known, whereas the curve itself is regarded as something to be derived. In other words, often an equation is given beforehand, and a curve is thereby determined; such an approach is dictated by the necessity of geometric representation of functional relations.

If an equation is given and we are to answer the question: What curve is represented by this equation? (or, what is the curve having this as its equation?), then it is convenient to use the definition phrased as follows:

The curve represented by a given equation (referred to some coordinate system) is the locus of all those points of the plane whose coordinates satisfy the equation.

36. The curve represented by an equation of the form $y = f(x)$ is called *the graph* of the function $f(x)$. It may also be said that the curve represented by an arbitrary equation $F(x, y) = 0$ is the graph of the functional relation between x and y , established by this equation.

37. Since the quantities x, y are regarded as the coordinates of a variable point, they are called the *current coordinates*. If, instead of cartesian coordinates, any other coordinate system is used as the reference system, then the current coordinates should be denoted by different letters, according to the notation adopted for the system used.

38. Let us consider a few elementary examples of curves represented by equations.

1. Given the equation $x - y = 0$. Rewriting it in the form $y = x$, we conclude that the equation is satisfied by the coordinates of those points, and those only, which are situated in the first or the third quadrant, at equal distances from the coordinate axes. Thus, the locus of points whose coordinates satisfy the given equation is the line bisecting the first and third quadrants (Fig. 23); this line is the curve represented by the equation $x - y = 0$ (also, this line is the graph of the function $y = x$).

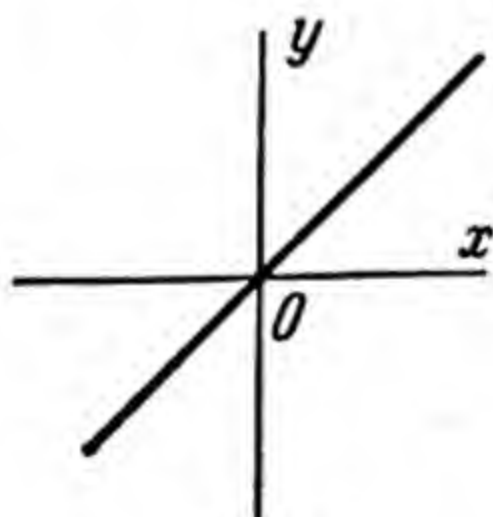


Fig. 23.

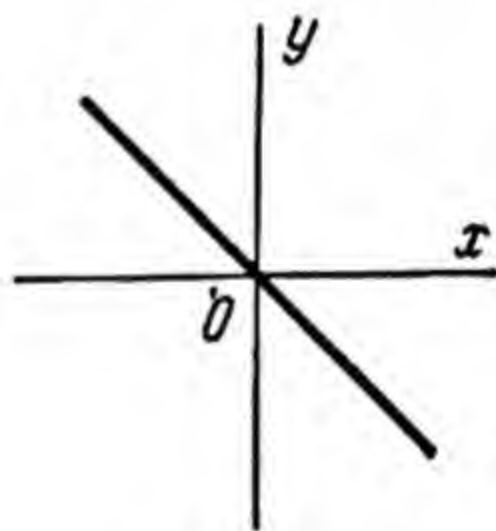


Fig. 24.

2. Given the equation $x + y = 0$. Rewriting it in the form $y = -x$, we conclude that this equation is satisfied by the coordinates of those points, and those only, which are equidistant from the coordinate axes and found in the second or the fourth quadrant. Thus, the locus of points whose coordinates satisfy the given equation is the line bisecting the second and fourth quadrants (Fig. 24); this line is the curve represented by the equation $x + y = 0$ (also, this line is the graph of the function $y = -x$).

3. Given the equation $x^2 - y^2 = 0$. Rewriting it in the form $(x - y)(x + y) = 0$, we conclude that this equation is satisfied

by the coordinates of those points, and those only, which satisfy *either* the equation $x - y = 0$, or the equation $x + y = 0$. Thus, the curve represented by the equation $x^2 - y^2 = 0$ consists of the points of the two lines bisecting the quadrants (Fig. 25).

4. Given the equation $x^2 + y^2 = 0$. Since, for real x and y , the numbers x^2 and y^2 cannot differ in sign, it follows that they cannot cancel out when added together; consequently, if $x^2 + y^2 = 0$, then $x = 0$ and $y = 0$. Thus, the given equation is satisfied by the coordinates of the point $O(0, 0)$ alone. This means that the locus of points whose coordinates satisfy the equation $x^2 + y^2 = 0$ consists of a single point. In this case, the equation is said to represent a *degenerate curve*.

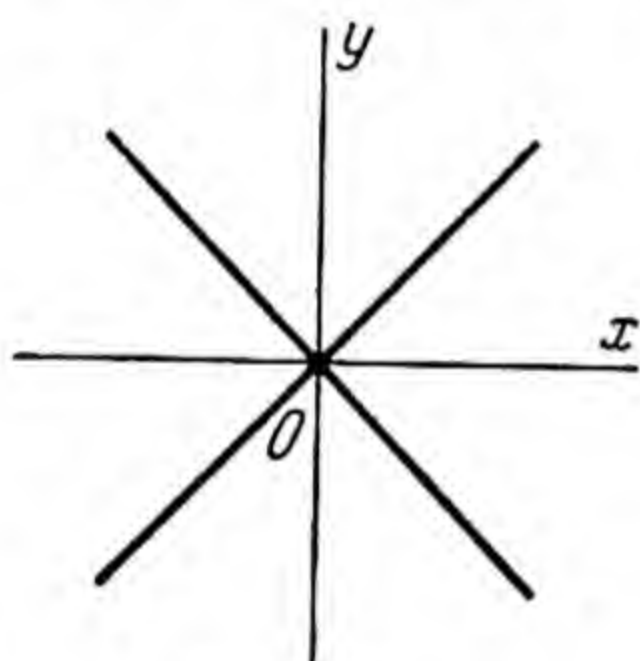


Fig. 25.

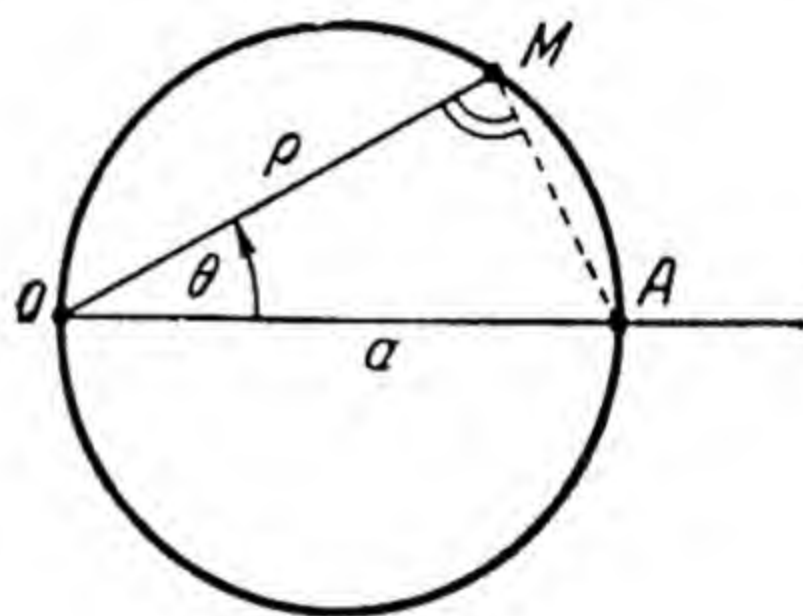


Fig. 26.

5. Let $x^2 + y^2 + 1 = 0$ be the given equation. Since, for any real x and y , the numbers x^2 and y^2 are non-negative, it follows that $x^2 + y^2 + 1 > 0$. Hence, the given equation cannot be satisfied by the coordinates of any point at all, and it does not represent any geometric object in the plane.

6. Consider the equation $\rho = a \cos \theta$, where a is a positive number, and the variables ρ and θ denote polar coordinates. Let M be a point with polar coordinates (ρ, θ) , and let A be the point with polar coordinates $(a, 0)$. If $\rho = a \cos \theta$, then $\angle OMA$ is a right angle, and conversely. Hence, the locus of points whose polar coordinates satisfy the equation $\rho = a \cos \theta$ is a circle of diameter OA (Fig. 26).

7. Consider the equation $\rho = a \theta$, where a is a positive number. To visualise the curve represented by this equation, let θ increase from zero, and observe the motion of a variable point $M(\rho, \theta)$, whose coordinates are related by the given equation. If $\theta = 0$, then also $\rho = 0$; as θ increases from zero, ρ increases in proportion to θ (the number a serving as the factor of proportionality). We note that the variable point $M(\rho, \theta)$, starting from

the pole of the chosen polar system of coordinates, moves about the pole (in the positive direction) and simultaneously recedes from the pole. Thus, the point M describes a kind of *spiral*; the spiral represented by the equation $\rho = a\theta$ is called the *spiral of Archimedes* (Fig. 27).

Each time the point $M(\rho, \theta)$, moving along the spiral of Archimedes from any initial position, makes a complete turn about the pole in the positive direction, the angle θ increases by the amount 2π , and the polar radius ρ , by $2a\pi$. It follows that the spiral of Archimedes cuts every polar ray into equal segments (excepting

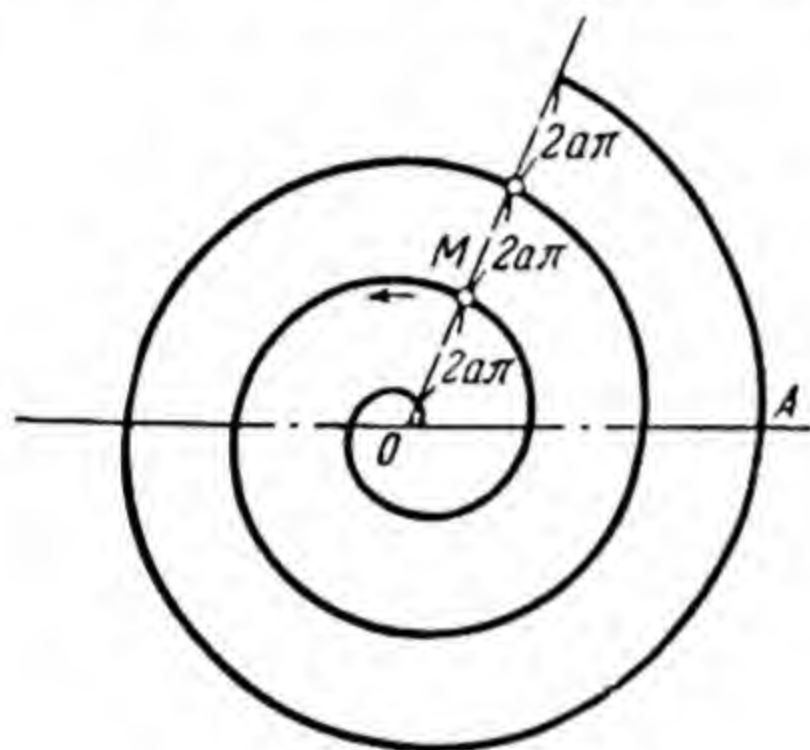


Fig. 27.

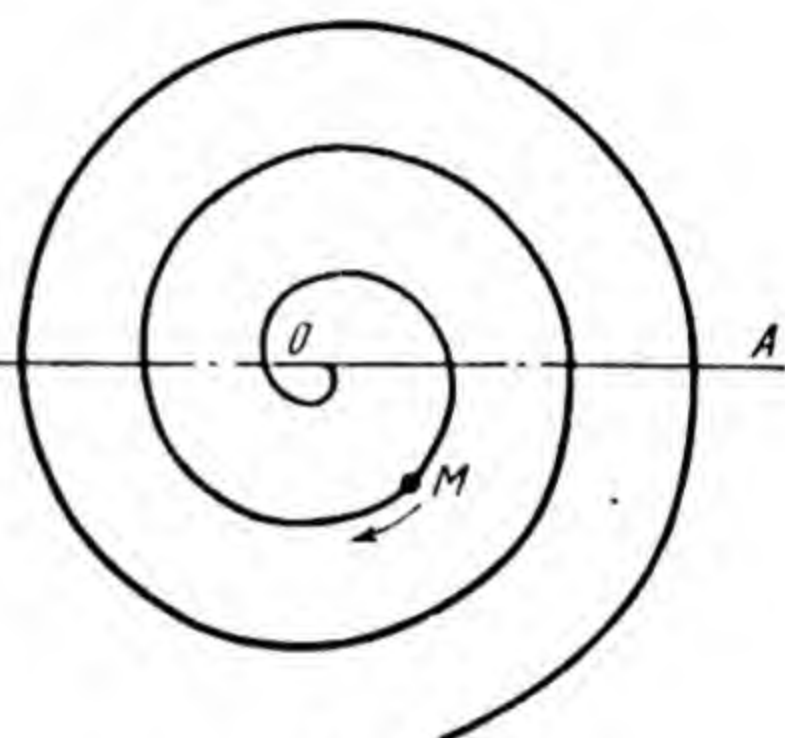


Fig. 28.

the segment adjacent to the pole); all these segments have a constant length of $2a\pi$.

The equation $\rho = a\theta$ in which a is a *negative* number represents the "inverted" spiral of Archimedes, whose points correspond to negative values of θ (Fig. 28).

8. Given the equation $\rho = \frac{a}{\theta}$, where a is a positive number; let us investigate the curve represented by this equation. Take any positive value of θ , say $\theta = \frac{\pi}{2}$; the corresponding point will be $M_1\left(\frac{2a}{\pi}, \frac{\pi}{2}\right)$. If now θ increases indefinitely, then ρ , being inversely proportional to θ , tends to zero; consequently, the variable point $M(\rho, \theta)$ moves around the pole in the positive direction and, at the same time, approaches indefinitely close to the pole (Fig. 29). Next, let θ decrease, starting from the value $\frac{\pi}{2}$ and tending to zero; then $\rho \rightarrow \infty$ and the point $M(\rho, \theta)$ recedes to infinity. To investigate the motion of the point M in greater detail, project the point M on the polar axis and denote the projec-

tion by P ; then, obviously, $PM = \rho \sin \theta$ (see the second of formulas (1), Art. 15). In virtue of the given equation, $\rho \sin \theta = a \frac{\sin \theta}{\theta}$. Now, it is known from the calculus that $\frac{\sin \theta}{\theta} \rightarrow 1$ as $\theta \rightarrow 0$.

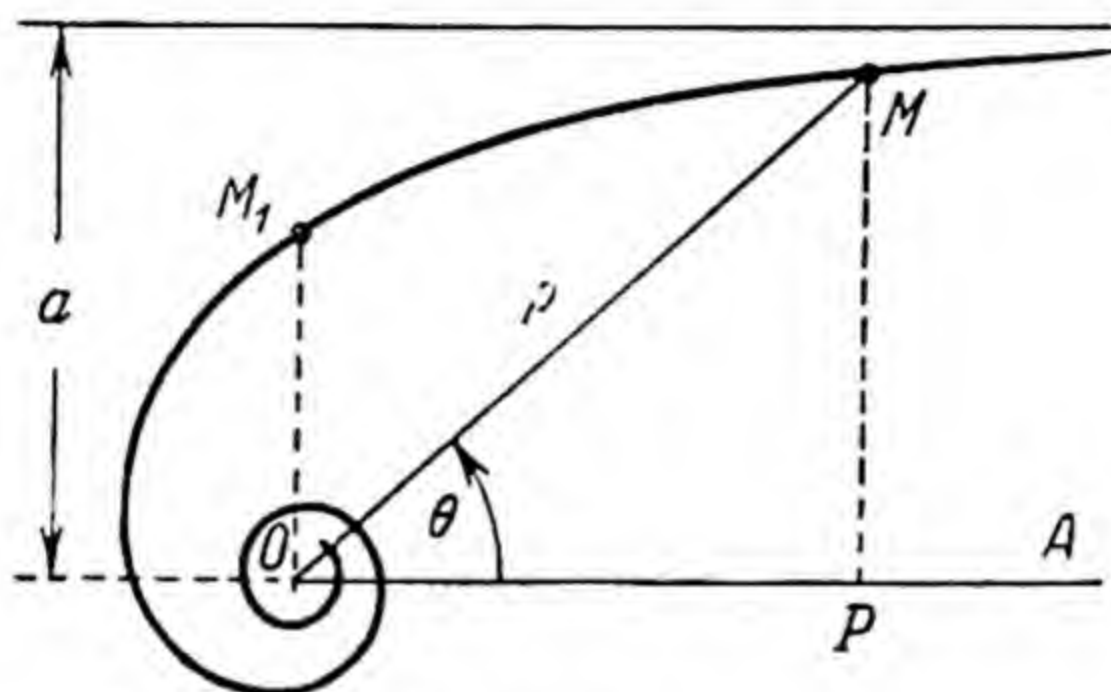


Fig. 29.

Consequently, as $\theta \rightarrow 0$, the value of PM tends to a . Hence we can conclude that, as the point M tends to infinity, it approaches the straight line which runs parallel to the polar axis at the distance a from the latter.

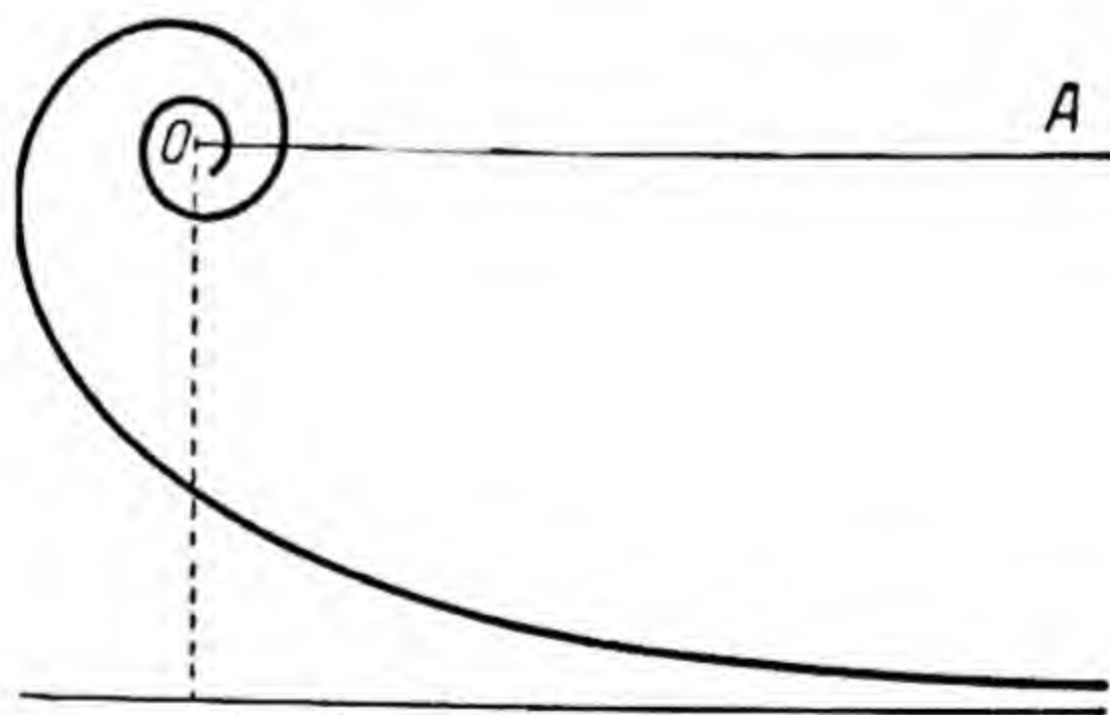


Fig. 30.

We see that the given equation, like that of the previous example, represents a spiral (in this case, the *hyperbolic spiral*).

The equation $\rho = \frac{a}{\theta}$ in which a is a *negative* number, represents the "inverted" hyperbolic spiral, whose points correspond to negative values of θ (Fig. 30).

9. Given the equation $\rho = a^\theta$, where a is a positive number greater than unity. This equation represents a spiral called *the logarithmic spiral*.

To visualise this particular kind of spiral, let $\theta \rightarrow +\infty$; then $\rho = a^\theta \rightarrow +\infty$, and hence a variable point $M(\rho, \theta)$, moving about the pole in the positive direction, recedes indefinitely from the pole. Each time the point M , starting from any position, makes a complete turn about the pole in the positive direction, 2π is added to the polar angle of M , while its polar radius is multiplied by $a^{2\pi}$ (since $a^{\theta+2\pi} = a^\theta a^{2\pi}$). Thus, with each turn about the pole,

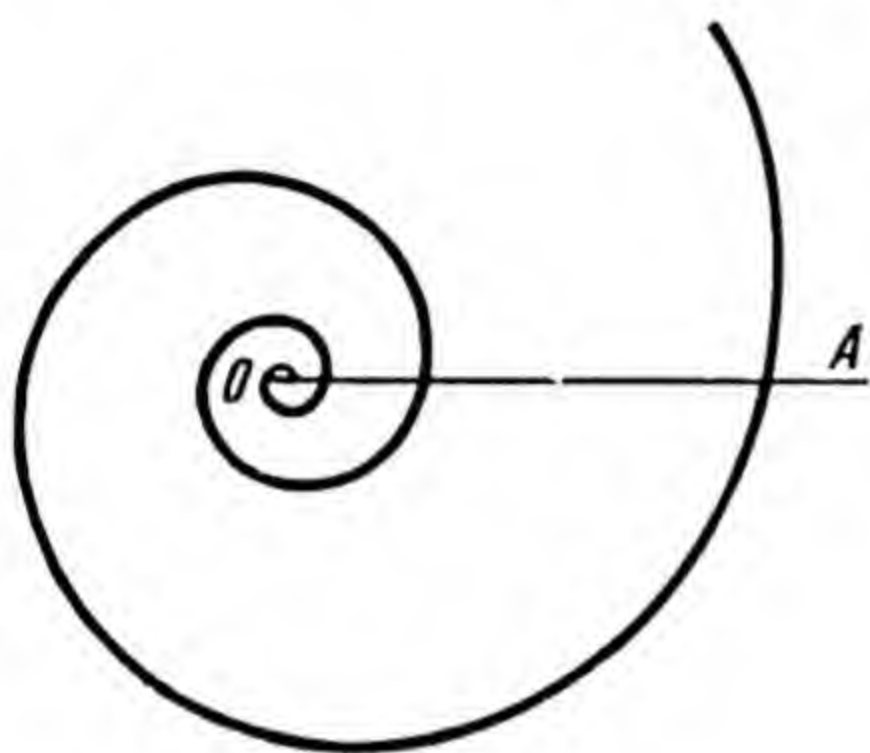


Fig. 31.

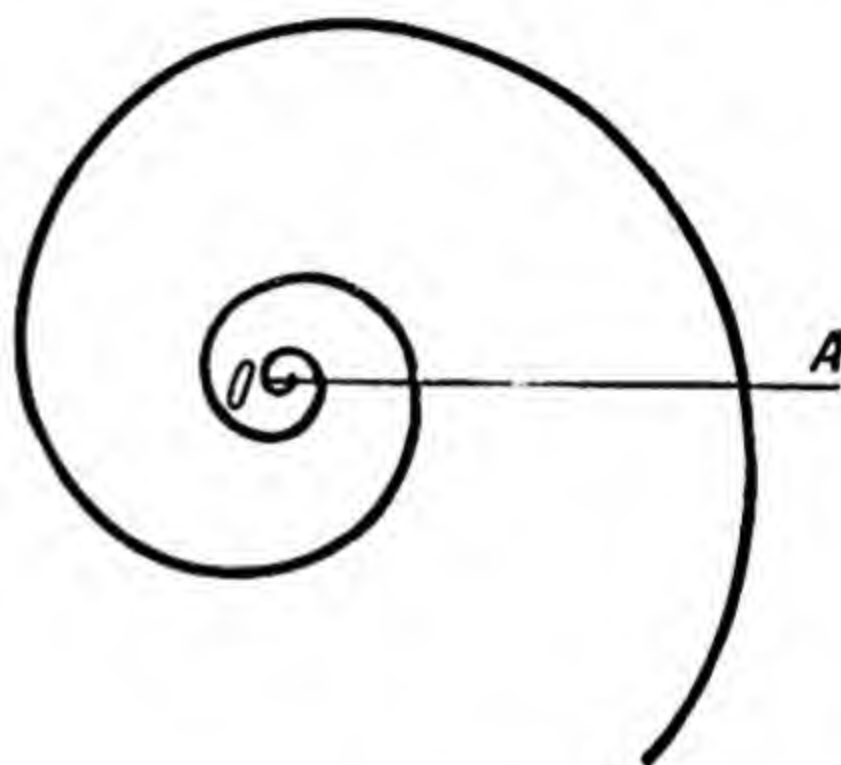


Fig. 32.

the polar radius of M increases in geometric progression ($a^{2\pi}$ being the common ratio of the progression).

Now let $\theta \rightarrow -\infty$; then $\rho \rightarrow 0$ and the point M , turning about the pole (in the negative direction) approaches it indefinitely near (Fig. 31).

If a is less than unity (and is still positive), the equation $\rho = a^\theta$ represents the "inverted" logarithmic spiral (Fig. 32). In this case, as the point M rotates about the pole in the positive direction, it approaches indefinitely close to the pole; as M rotates in the negative direction, it recedes from the pole without limit (since, for the case $0 < a < 1$, we have $a^\theta \rightarrow 0$ as $\theta \rightarrow +\infty$, and $a^\theta \rightarrow +\infty$ as $\theta \rightarrow -\infty$).

If $a = 1$, the equation $\rho = a^\theta$ represents a circle, because $\rho = 1$ for any θ .

In the above examples, the equations were of such simple form as to permit an immediate visualisation of the associated curves. In more complicated cases, even an approximate tracing

(to a preassigned degree of accuracy) of the curve in question may present great difficulties and require the use of various methods of the calculus and analytic geometry.

§ 12. Examples of Deriving the Equation of a Given Curve

39. In the preceding section, we discussed several examples dealing with the determination of curves from the given equations. We shall now consider some examples of an opposite character; in each of these examples, the curve is defined in purely geometric terms, and it is required to find (to "derive") the equation of the curve from this geometric definition.

If a curve has been defined as the locus of points satisfying a certain condition, then, by expressing this condition in terms of coordinates, we shall obtain a definite relation between the coordinates. This will be the equation of the given curve, for it will be satisfied by the coordinates of a point if, and only if, the position of the point obeys the imposed condition, that is, if the point lies on the given curve.

40. **Example.** Given a rectangular cartesian coordinate system, and let $C(\alpha, \beta)$ be the centre of a circle whose radius is equal to r (Fig. 33). Derive the equation of the circle.

Let M be a variable point, and denote its coordinates (that is, the current coordinates) by the letters x, y . The given circle is the locus of points each of which is at a distance r from the point C ; hence, the point M lies on the given circle if, and only if,

$$CM = r. \quad (1)$$

By formula (2) of Art. 18, $CM = \sqrt{(x - \alpha)^2 + (y - \beta)^2}$. Substituting this expression for CM in (1), we have

$$\sqrt{(x - \alpha)^2 + (y - \beta)^2} = r. \quad (2)$$

We have thus found the equation that connects the variables x, y and is satisfied by the coordinates of the points, and those only, which lie on the given circle. Consequently, this is the required equation. Our problem is solved.

41. Squaring both sides of (2), we obtain the standard form of the equation of a circle with centre $C(\alpha, \beta)$ and radius r :

$$(x - \alpha)^2 + (y - \beta)^2 = r^2. \quad (3)$$

* See the footnote on page 13.

Equation (3) appears in many geometric problems^{*)}. Setting $\alpha = 0$, $\beta = 0$ in this equation, we obtain the equation of a circle with centre at the origin:

$$x^2 + y^2 = r^2. \quad (4)$$

42. Example. Derive the equation of the path traced by a point M , whose distance from the point $B(8, 0)$ is twice its distance from the point $A(2, 0)$ at each instant of the motion.

Let x, y be the coordinates of the point M (that is, the current coordinates). By hypothesis, the point M is always two times closer to A than to B ; that is,

$$2AM = BM. \quad (5)$$

By formula (2) of Art. 18,

$$AM = \sqrt{(x-2)^2 + y^2}, \quad BM = \sqrt{(x-8)^2 + y^2}.$$

Hence, from (5), we have

$$2\sqrt{(x-2)^2 + y^2} = \sqrt{(x-8)^2 + y^2}. \quad (6)$$

We have obtained an equation connecting the variables x and y . It is satisfied by the coordinates of all points of the path considered and by the coordinates of no other point in the plane. Consequently, this is the required equation, and the problem is solved. It remains only to reduce the equation to a more convenient form, which is achieved as follows. Squaring both sides of (6), we obtain the equation

$$4[(x-2)^2 + y^2] = (x-8)^2 + y^2,$$

equivalent to equation (6)^{**)} . Removing the parentheses, we find

$$4x^2 - 16x + 16 + 4y^2 = x^2 - 16x + 64 + y^2,$$

or

$$x^2 + y^2 = 16.$$

With the equation of the path put in this form, the path can readily be visualised.

For, comparing the derived equation with equation (4) of Art. 41, we conclude that the path under discussion is a circle with centre at the origin and with radius $r = 4$.

43. Example. Find the equation of a straight line in polar coordinates, given that ρ is the distance from the pole to the straight line, and θ_0 the angle from the polar axis to the ray drawn from the pole perpendicular to the line (Fig. 34).

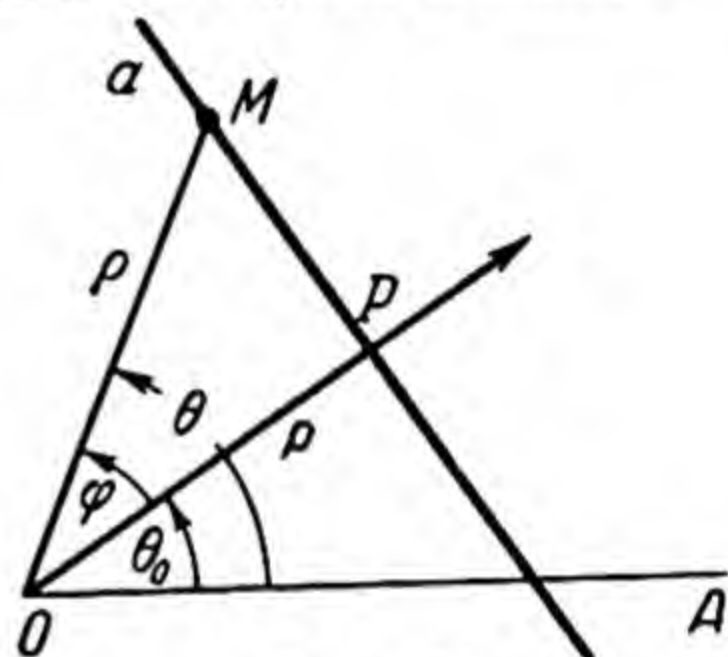


Fig. 34.

^{*)} Squaring both sides of an equation may result in an equation which will not be equivalent to the original equation; that is, the resulting equation may be satisfied by such values of x and y as will not satisfy the original equation. In the present case, however, this is not so: equations (2) and (3) are equivalent. In fact, extracting the square root of both sides of (3) yields $\pm \sqrt{(x-\alpha)^2 + (y-\beta)^2} = \pm r$. But the right-hand member must be taken here with the plus sign, or else the equation will not be true. Thus, not only equation (3) follows from equation (2), but also (2) follows from (3).

^{**)} This is proved in the same way as the analogous statement in Art. 41 (see the preceding footnote).

Solution. Let a be an arbitrary straight line, and p the foot of the perpendicular dropped upon it from the pole O ; by hypothesis, $OP = p$ and the angle θ_0 from the polar axis OA to the ray OP are known. Let us consider an arbitrary point $M(\rho, \theta)$ in the plane. Obviously, the point M lies on the line a if, and only if, the projection of the point M on the ray OP coincides with the point P , that is, if $\rho \cos \varphi = p$, where $\varphi = \angle POM$. Noting that $\varphi = \theta - \theta_0$ (or $\varphi = \theta_0 - \theta$), we hence obtain $\rho \cos (\theta - \theta_0) = p$ as the required equation of the line a ; ρ and θ are here the current coordinates.

§ 13. The Problem of the Intersection of Two Curves

44. The following problem has often to be solved in analytic geometry:

Given the equations of two curves,

$$F(x, y) = 0, \quad \Phi(x, y) = 0,$$

to find their points of intersection.

As always in analytic geometry, "to find the points" means here "to calculate their coordinates". The principle underlying the solution of this problem becomes immediately clear from the definition of the equation of a curve (Art. 34). For, each point of intersection of the given curves is common to the two curves. Consequently, the coordinates of such a point must satisfy both the equation $F(x, y) = 0$ and the equation $\Phi(x, y) = 0$. All the points of intersection of these curves will be found by solving the equations simultaneously, each solution of the system

$$\left. \begin{aligned} F(x, y) &= 0, \\ \Phi(x, y) &= 0 \end{aligned} \right\}$$

determining one of the required points. Of course, computational difficulties may be encountered in practical application of this general principle.

Example 1. Given the equations of two circles: $(x - 1)^2 + (y - 3)^2 = 4$ and $(x - 3)^2 + (y - 5)^2 = 4$. Find their points of intersection.

Solution. By removing the parentheses and transposing all terms to the left-hand side, the equations may be reduced to the form

$$x^2 + y^2 - 2x - 6y + 6 = 0, \quad x^2 + y^2 - 6x - 10y + 30 = 0. \quad (1)$$

Subtracting the second equation from the first, we get $4x + 4y - 24 = 0$, or $y = -x + 6$. Together with the first of the given equations, this last equation forms the system

$$\left. \begin{aligned} x^2 + y^2 - 2x - 6y + 6 &= 0, \\ y &= -x + 6. \end{aligned} \right\} \quad (2)$$

The systems (1) and (2) are equivalent*). Accordingly, in order to solve our problem, we have merely to solve the system (2). Substitution of $y = -x + 6$ in the first of equations (2) yields $x^2 + x^2 - 12x + 36 - 2x + 6x - 36 + 6 = 0$, or $x^2 - 4x + 3 = 0$. Hence, $x_{1,2} = 2 \pm \sqrt{4-3}$; that is, $x_1 = 1$, $x_2 = 3$. The values of y , corresponding to these values of x , are found from the equation $y = -x + 6$; we have $y_1 = 5$ for $x_1 = 1$, and $y_2 = 3$ for $x_2 = 3$. Thus, the required points are (1, 5) and (3, 3).

Example 2. Given the equations of two curves: $x + y = 0$ (the bisector of the second quadrant), and $(x-5)^2 + y^2 = 1$ (a circle). Find their points of intersection.

Solution. We form the system

$$\begin{cases} (x-5)^2 + y^2 = 1, \\ x + y = 0. \end{cases}$$

From the second of these equations, $y = -x$. Substituting this in the first equation, we obtain $(x-5)^2 + x^2 = 1$, or $x^2 - 5x + 12 = 0$. Hence,

$$x_{1,2} = \frac{5}{2} \pm \sqrt{\frac{25}{4} - 12} = \frac{5}{2} \pm \frac{\sqrt{-23}}{2}.$$

Since $\sqrt{-23}$ is an imaginary number, we conclude that the system has no real solutions and, consequently, the given curves do not intersect.

§ 14. Parametric Equations of a Curve

45. Let there be chosen a coordinate system, and let

$$\begin{cases} x = \varphi(t), \\ y = \psi(t) \end{cases} \quad (1)$$

be two given functions of a single variable t .

We shall agree to regard the quantities x and y , for every value of t , as the coordinates of a point M . The quantities x and y , in general, change with t ; it follows that the point M moves in the plane. Relations (1) are called the *parametric equations of the path traced by the point M* ; the independent variable t is called a *variable parameter*.

Parametric equations play an important role in mechanics, where they are used as the so-called *equations of motion*. For, if a particle M moves in a plane, it has definite coordinates x , y at any instant of time t . The equations expressing x and y as functions of time t are called the equations of motion of the point M ; they are of the form (1).

In mechanics, the motion of a particle is regarded as mathematically determined if its equations have been found.

* Since the system (2) has been derived from the system (1), which is, in its turn, readily derivable from the system (2),



46. Let $x = \varphi(t)$, $y = \psi(t)$ be the parametric equations representing some curve as the path of a point $M(x, y)$.

If $F(x, y) = 0$ is a consequence of the given equations, then it is satisfied by the coordinates $x = \varphi(t)$, $y = \psi(t)$ of the point M for every t . Hence the point M moves along the curve $F(x, y) = 0$. If, in so doing, the point M traces the entire curve, then $F(x, y) = 0$ is an ordinary equation of the path of M . The derivation of $F(x, y) = 0$ as a consequence of the parametric equations $x = \varphi(t)$, $y = \psi(t)$ is called *the elimination of the parameter*.

Example. The equations $x = r \cos t$, $y = r \sin t$ are the parametric equations of a circle with centre at the origin and with radius r . For, squaring these equations and adding them, term by term, we obtain $x^2 + y^2 = r^2$ as their consequence. It is hence evident that the point $M(x, y)$ moves on this circle. Moreover, since the parameter t can assume all possible values, the ray OM (which makes the angle t with the axis Ox) can occupy all possible positions. Consequently, the point M traces the entire circle (doing this an indefinite number of times as t increases indefinitely).

47. Let $\rho = f(\theta)$ be the polar equation of a certain curve. The same curve can be represented in cartesian coordinates by the parametric equations

$$\begin{aligned}x &= f(\theta) \cos \theta, \\y &= f(\theta) \sin \theta.\end{aligned}$$

These equations are obtained by simply substituting $f(\theta)$ for ρ in the formulas $x = \rho \cos \theta$, $y = \rho \sin \theta$ (see Art. 15).

Example. The polar equations of the spiral of Archimedes, the hyperbolic spiral, and the logarithmic spiral are $\rho = a\theta$, $\rho = \frac{a}{\theta}$, and $\rho = a^\theta$, respectively (see Art. 38). Hence, their parametric equations in cartesian coordinates are:

$$x = a\theta \cos \theta, \quad y = a\theta \sin \theta$$

for the spiral of Archimedes;

$$x = \frac{a \cos \theta}{\theta}, \quad y = \frac{a \sin \theta}{\theta}$$

for the hyperbolic spiral;

$$x = a^\theta \cos \theta, \quad y = a^\theta \sin \theta$$

for the logarithmic spiral.

In all these cases, the polar angle θ of the variable point serves as the parameter.

§ 15. Algebraic Curves

48. Analytic geometry has as its main subject of study the curves represented, in rectangular cartesian coordinates, by *algebraic equations*. These are equations of the following forms:

$$Ax + By + C = 0; \quad (1)$$

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0; \quad (2)$$

$$Ax^3 + Bx^2y + Cxy^2 + Dy^3 + Ex^2 + Fxy + Gy^2 + Hx + Iy + K = 0; \quad (3)$$

.

A, B, C, D, E , etc., denote here fixed numbers and are called the coefficients of these equations.

Equation (1) is called *the general equation of the first degree* (its coefficients may have any values whatsoever, provided only that the equation does contain terms of the first degree; that is, A and B cannot both be zero at the same time); equation (2) is called *the general equation of the second degree* (its coefficients may have any values, provided only that the equation does contain terms of the second degree, which means that the three coefficients A, B, C cannot all be zero at the same time); equation (3) is called *the general equation of the third degree* (its coefficients may have any values, provided only that the four coefficients A, B, C, D are not all simultaneously zero). The equations of the fourth, fifth, etc., degrees have analogous forms.

Examples of *non-algebraic* equations are

$$y - \sin x = 0,$$

$$y - \log x = 0,$$

$$y - 10^x = 0,$$

$$10^x - 5^y + 1 = 0,$$

$$2^{xy} - x - y = 0.$$

A curve represented, in a rectangular cartesian system of coordinates, by an algebraic equation of degree n is called an algebraic curve of the n th order.

49. Theorem 8. *A curve represented by an algebraic equation of degree n in rectangular cartesian coordinate system, will be represented in any other rectangular cartesian system by another algebraic equation of the same degree n .*

Proof. Let some curve be represented by an algebraic equation of degree n in a coordinate system with axes Ox and Oy . When replacing this system by another rectangular system with axes $O'x'$, $O'y'$, the coordinates of all points of the plane are transformed according to formulas of the form

$$\left. \begin{aligned} x &= x' \cos \alpha \mp y' \sin \alpha + a, \\ y &= x' \sin \alpha \pm y' \cos \alpha + b, \end{aligned} \right\} \quad (4)$$

where the signs before the second terms of the right-hand members are chosen in conformity with the note made at the end of Art. 29. In order to obtain the equation of the same curve in the new coordinates, we must replace the variables in its original equation according to formulas (4). The left-hand

member of the original equation is a sum of monomials, each of which is a product (taken with some coefficient) of non-negative whole powers of the variables x and y . On replacing x and y according to formulas (4) and removing all parentheses, we shall obtain, on the left-hand side of the transformed equation, a sum of new monomials, each of which will be a product (taken with some coefficient) of non-negative whole powers of the new variables x' and y' . Consequently, the *algebraic* character of the equation is *preserved* under such a transformation.

Next, we must prove that the *degree* of the equation remains *unchanged*. This is almost obvious. For, since formulas (4) are first-degree equations in x' and y' , the replacement of x and y according to these formulas and the removal of all parentheses in the left-hand member of the transformed equation cannot result in the appearance of any monomial of a degree^{*)} higher than the n th with respect to the new variables x' and y' .

Hence, the degree of an algebraic equation cannot be raised by any such transformation. It remains, however, to be shown that the degree of the equation cannot be lowered by any such transformation (i. e., that the highest terms cannot all cancel out after the transformation). But if a transformation from one rectangular cartesian system of coordinates to another such system could result in lowering the degree of an algebraic equation, then the inverse transformation would necessarily raise its degree, and this we have just shown to be impossible. The proof is thus complete.

The theorem just proved shows that the algebraic character and degree of an equation are *intrinsic properties of the associated algebraic curve itself*, which means that they are *independent of the choice of coordinate axes*.

The general theory of algebraic curves forms the subject of special treatises on analytic geometry. The present book deals systematically with curves of the first and the second order only.

In the next few sections, it will be established that curves of the first order are straight lines (and only straight lines).

^{*)} The degree of a monomial is defined as the sum of the exponents of the variables contained in the monomial.

CURVES OF THE FIRST ORDER

§ 16. The Slope of a Straight Line

50. Let there be given a rectangular cartesian system of coordinates and a straight line. Denote by α the angle through which the axis Ox must be turned to reach coincidence with one of the two directions of the given straight line; this angle will be taken with a plus or minus sign according as the turning is in the positive or the negative direction. We shall call α the angle of inclination of the given straight line (with respect to the axis Ox).

If, by turning the axis Ox through a certain angle, Ox is made to coincide with one of the directions of the given straight line,

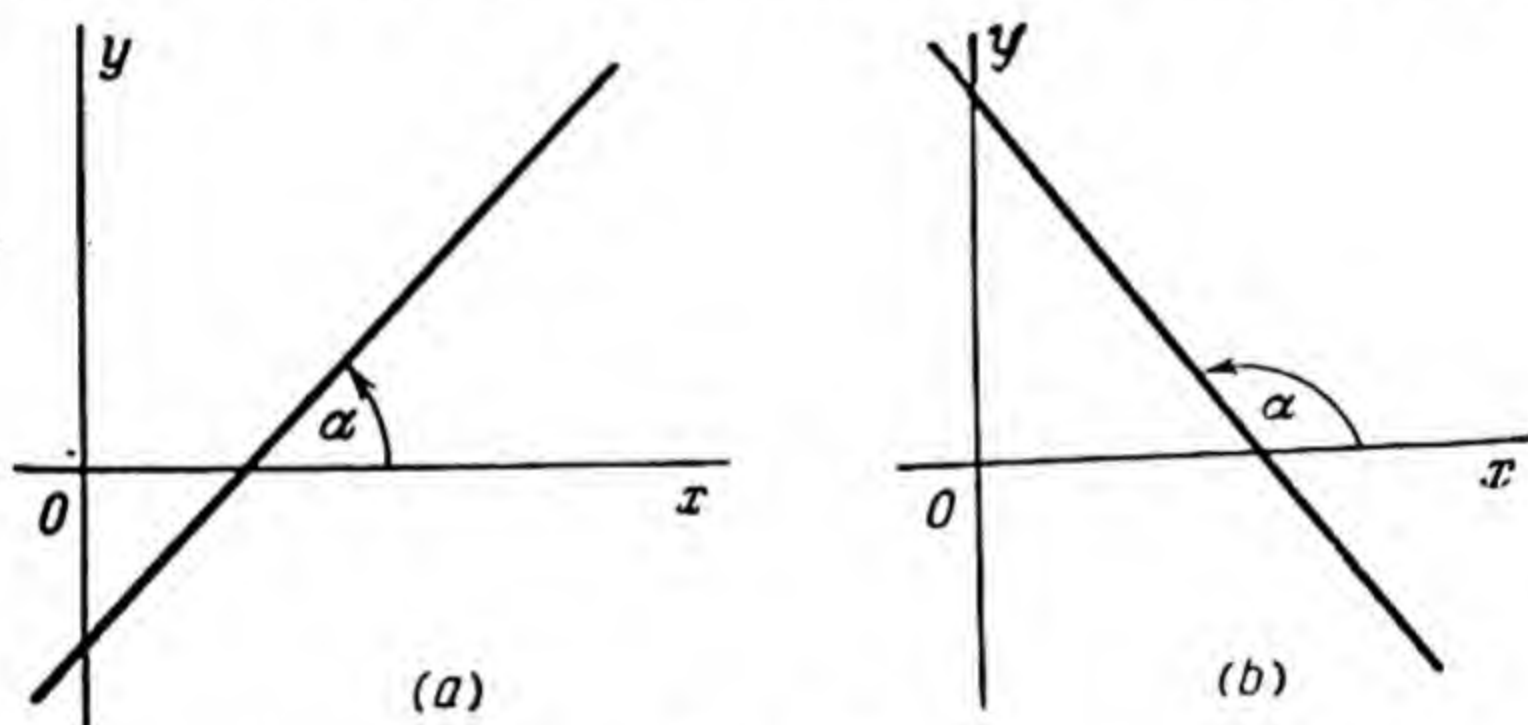


Fig. 35.

then each additional rotation through $\pm\pi$, or $\pm2\pi$, or $\pm3\pi$, etc., will again bring the axis into coincidence with one of the directions of the given line. Thus, the angle α can have an infinity of values differing from one another by the amount $\pm n\pi$, where n is a natural number. For the most part, it is the smallest positive value of the angle α (Fig. 35 a, b) that is taken as the angle of inclination of a line; in the case of a line parallel to the axis Ox , its angle of inclination is considered equal to zero.

It is important to note that, for a given line, all values of its angle of inclination have the same trigonometric tangent, since $\tan(\alpha \pm n\pi) = \tan \alpha$.

51. The tangent of the angle of inclination of a straight line is called the slope of that line.

Denoting the slope by the letter k , we can write the above definition symbolically:

$$k = \tan \alpha. \quad (1)$$

In particular, if $\alpha = 0$, then also $k = 0$, which means that the slope of a line parallel to the axis Ox is equal to zero. If $\alpha = \frac{\pi}{2}$, then $k = \tan \alpha$ has no arithmetical meaning (is represented by no number), that is, the slope of a line perpendicular to the axis Ox fails to exist. However, it is very often said that, if a straight

line is perpendicular to the axis Ox , its slope "becomes infinite", thereby expressing the fact that, as $\alpha \rightarrow \frac{\pi}{2}$, $k \rightarrow \infty$.

The slope of a straight line is the essential characteristic of the direction of that line and is constantly used in analytic geometry and its applications.

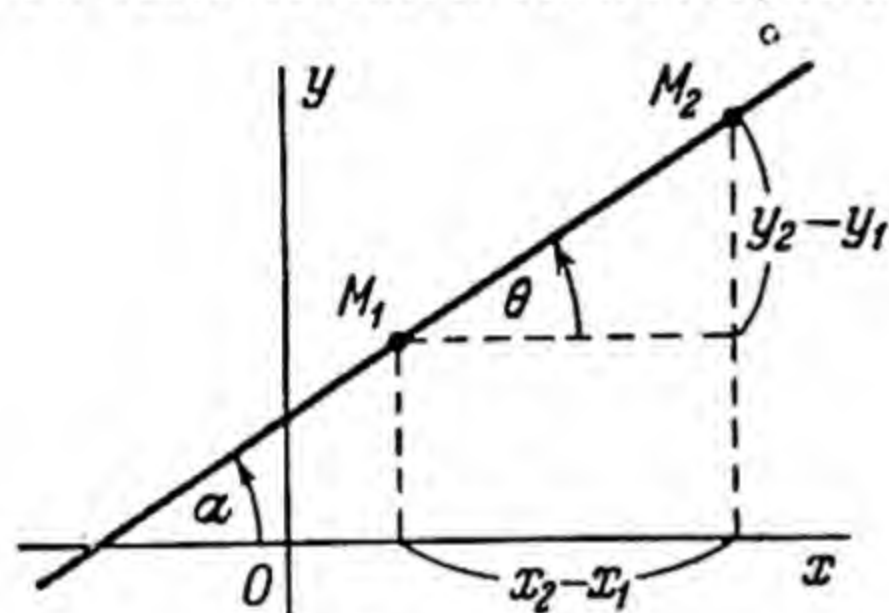


Fig. 36.

52. Let there be given an arbitrary straight line, provided only that it is not perpendicular to the axis Ox . Take any two points $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$ on this line. The polar angle θ of the segment $\overline{M_1M_2}$ is equal to the angle of inclination of the given line, so that the tangent of the angle θ is equal to the slope of this line (Fig. 36); hence, by formula (6) of Art. 19, we have

$$k = \frac{y_2 - y_1}{x_2 - x_1} \quad (2)$$

(this relation is also apparent from Fig. 36). Formula (2) expresses the slope of the straight line passing through two given points.

§ 17. The Slope-intercept Equation of a Straight Line

53. Let there be given an arbitrary straight line, provided that, as before, it is not perpendicular to the axis Ox . We shall now derive the equation of this line, assuming that we know its slope k and its y -intercept b (that is, the value b of the directed segment \overline{OB} cut off by this line on the axis Oy ; see Fig. 37).

Let M be a variable point, with x, y as its (current) coordinates, and consider also the point $B (0, b)$ at which the line cuts the axis Oy . Let us compute the value of the right member in formula (2) of Art. 52, taking the point B as M_1 and the point M as M_2 . If the point M lies on the given line, the computation will yield the slope of the line, that is,

$$\frac{y-b}{x} = k; \quad (3)$$

on the other hand, if M does not lie on the given line, this relation will not be valid. Consequently, (3) is the equation of the given straight line (this is also apparent from Fig. 37, when taking into account that $k = \tan \alpha$). Clearing of fractions and transposing b to the right side, we get

$$y = kx + b. \quad (4)$$

54. Thus, every straight line not perpendicular to the axis Ox can be represented by an equation of the form (4).

Conversely, every equation of the form (4) represents a straight line having slope k and y -intercept b . For, if $y = kx + b$ is the given equation, it is always possible to draw the line with the given slope k and making the given intercept b on the axis Oy , no matter what the numbers k and b ; but, according to the foregoing, the given equation will then be the equation of the line drawn. An equation of the form (4) is called the slope-intercept equation of a straight line.

Example. Draw the line whose equation is

$$y = \frac{3}{4}x + 2.$$

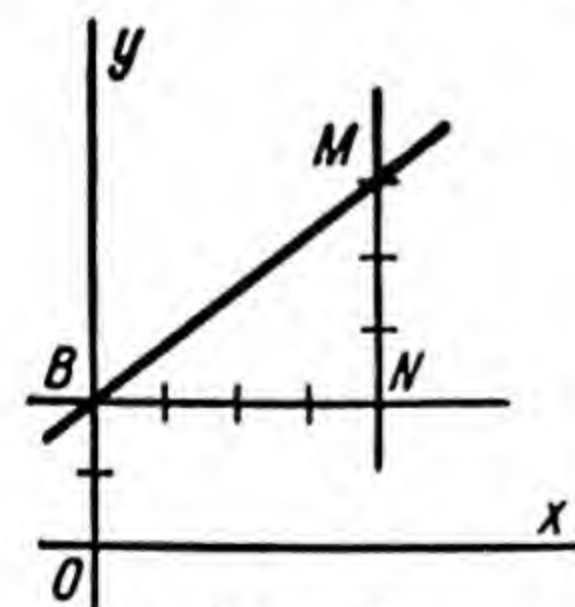


Fig. 38.

Solution. Mark off the segment $OB = 2$ on the axis Oy (Fig. 38); through the point B , draw the segment $BN = 4$ in the "right" direction, parallel to the axis Ox ; and lay off the segment $NM = 3$ from the point N "upwards" (in the direction of the axis Oy). Joining the points B and M will then give the required line (which cuts off the intercept $b = 2$ on Oy and makes with Ox an angle whose tangent is $\frac{3}{4}$).

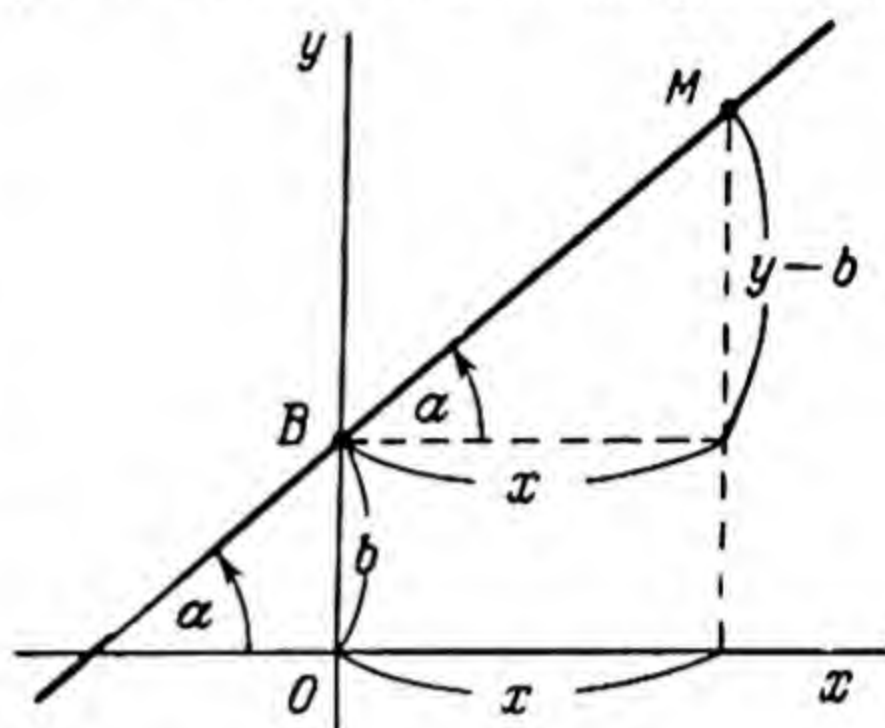


Fig. 37.

55. The function

$$y = kx + b$$

is called *linear*. From the foregoing, we can say that *the graph of a linear function is a straight line*.

When $b = 0$, we have

$$y = kx. \quad (5)$$

The variables x and y , connected by such a relation, are said to be *proportional*; the number k is called *the factor of proportionality*. From the above discussion, it is clear that *the graph of the function $y = kx$ is the line of slope k passing through the origin*.

56. In many cases it is necessary to find the equation of the straight line with given slope k and passing through the given point $M_1 (x_1, y_1)$. The desired equation is obtained directly from formula (2) of Art. 52. Let M be a variable point with (current) coordinates x, y . If M lies on the line of slope k passing through the point M , then, by virtue of formula (2) of Art. 52,

$$\frac{y - y_1}{x - x_1} = k; \quad (6)$$

if the point M does not lie on the line, relation (6) is not valid. Accordingly, (6) is the required equation, usually written in the form

$$y - y_1 = k(x - x_1). \quad (7)$$

Note. In the particular case when the point $B (0, b)$ is taken as $M_1 (x_1, y_1)$, equation (7) assumes the form (4).

57. By using relation (7), the following problem can easily be solved: *to find the equation of the line passing through the two given points $M_1 (x_1, y_1)$ and $M_2 (x_2, y_2)$* .

Using formula (2) of Art. 52, we find the slope of the line:

$$k = \frac{y_2 - y_1}{x_2 - x_1},$$

after which, from (7), we obtain the required equation

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1} (x - x_1).$$

It is customary to write this equation in the form

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1}. \quad (8)$$

Example. Find the equation of the line through the points $M_1 (3, 1)$ and $M_2 (5, 4)$.

Solution. Inserting the given coordinates in (8), we obtain

$$\frac{x-3}{2} = \frac{y-1}{3},$$

or $3x - 2y - 7 = 0$.

§ 18. Calculation of the Angle Between Two Straight Lines. Conditions for the Parallelism and Perpendicularity of Two Straight Lines

58. The problem of determining the angle between two straight lines is one of the standard problems of analytic geometry. We shall now derive a formula for calculating the angle between straight lines whose slopes are known (we assume that neither of the lines is perpendicular to the axis Ox).

Consider two straight lines; one of them (no matter which) will be referred to as the first line, and the other as the second line (see Fig. 39). Let k_1 and k_2 denote the respective slopes of these lines, and let φ be the angle which the second line makes with the first, that is, the angle through which the first line has to be turned to reach coincidence with

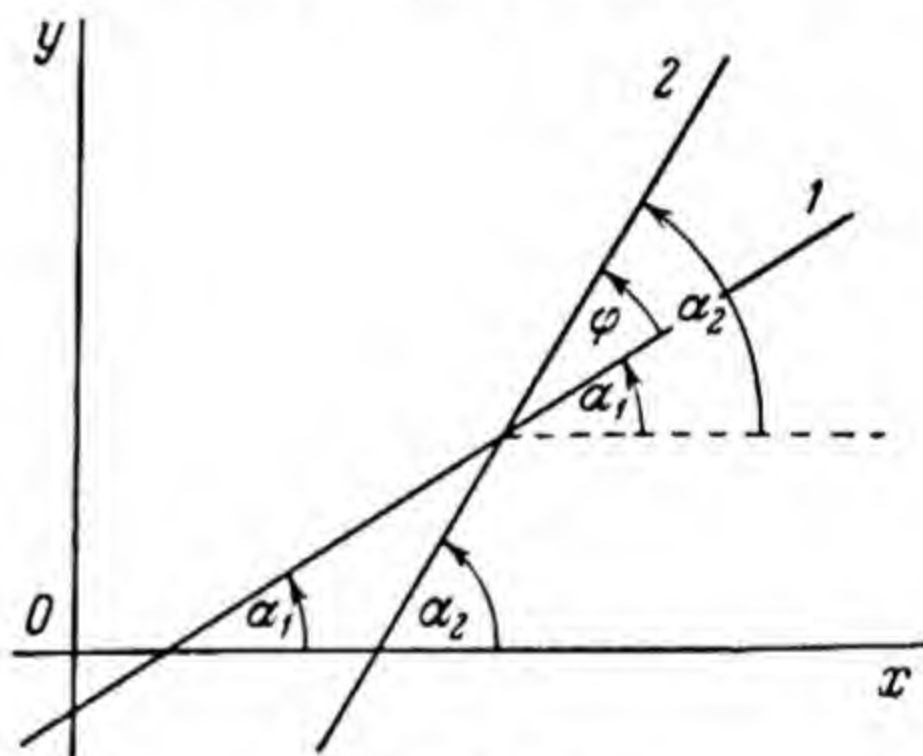


Fig. 39.

one of the directions of the second line. The angle φ will be taken with a plus or minus sign according as the turning is in the positive or the negative direction. It is this angle φ that we shall mean when speaking of the angle between two straight lines.

Let α_1 be the angle of inclination of the first line. If the axis Ox is turned through the angle α_1 , the axis will coincide with one of the directions of the first line; if the axis Ox is then given an additional turn through the angle φ , it will coincide with one of the directions of the second line. Thus, by adding the angle φ to the angle α_1 , the angle of inclination of the second line is obtained; we shall designate this angle of inclination as α_2 . Accordingly, we have $\alpha_1 + \varphi = \alpha_2$, or $\varphi = \alpha_2 - \alpha_1$. Hence

$$\tan \varphi = \tan (\alpha_2 - \alpha_1) = \frac{\tan \alpha_2 - \tan \alpha_1}{1 + \tan \alpha_1 \tan \alpha_2}.$$

Now, $\tan \alpha_1 = k_1$, $\tan \alpha_2 = k_2$; therefore,

$$\tan \varphi = \frac{k_2 - k_1}{1 + k_1 k_2}. \quad (1)$$

This is the formula we have been seeking.

If $\varphi = \frac{\pi}{2}$, the tangent of the angle φ has no arithmetical meaning ("becomes infinite"); then, and only then, the denominator of the right-hand member in (1) is equal to zero.

Example. Given the lines $y = -\frac{1}{7}x + 2$, $y = \frac{3}{4}x + 3$. Find the angle between them.

Solution. By formula (1),

$$\tan \varphi = \frac{\frac{3}{4} - \left(-\frac{1}{7}\right)}{1 + \frac{3}{4} \cdot \left(-\frac{1}{7}\right)} = \frac{21 + 4}{28 - 3} = 1.$$

Thus, one of the angles made by the given lines is equal to 45° .

59. In solving various problems of analytic geometry, it is often important to ascertain whether two straight lines, whose equations are known, happen to be parallel, or perpendicular, or neither. This question can also be easily resolved.

Let k_1 and k_2 be the known slopes of two straight lines whose angles of inclination will be denoted by α_1 and α_2 , respectively. Obviously, the given lines will be parallel if, and only if, their angles of inclination are equal, that is, if $\tan \alpha_1 = \tan \alpha_2$. But $\tan \alpha_1 = k_1$, $\tan \alpha_2 = k_2$. Hence, *the condition for the parallelism of two straight lines is the equality of their slopes:*

$$k_2 = k_1.$$

The given straight lines will be perpendicular if, and only if, the angle φ between them is equal to $\frac{\pi}{2}$, that is, if $\tan \varphi$ has no arithmetical meaning; in this case, the denominator of the right-hand member in formula (1) will become zero, so that we shall have $1 + k_1 k_2 = 0$. Consequently, *the condition for the perpendicularity of two straight lines is expressed by the relation*

$$k_1 k_2 = -1.$$

This last relation is usually written as

$$k_2 = -\frac{1}{k_1} \quad (2)$$

and, accordingly, the condition that two straight lines should be perpendicular is formulated as follows: *The slopes of perpendicular lines are negative reciprocals.*

Applying the formulas just derived, we can say at once that, for instance, the lines $y = \frac{2}{3}x + 1$, $y = \frac{2}{3}x + 5$ are parallel, whereas the lines $y = \frac{3}{4}x + 2$, $y = -\frac{4}{3}x + 3$ are mutually perpendicular.

Example. Find the projection of the point $P(4, 9)$ on the line passing through the points $A(3, 1)$ and $B(5, 2)$.

Solution. The required point will be found by solving simultaneously the equation of the line AB and the equation of the perpendicular dropped to this line from the point P . We shall begin by determining the equation of the line AB ; using relation (8) of Art. 57, we obtain

$$\frac{x-3}{2} = \frac{y-1}{1},$$

or

$$y = \frac{1}{2}x - \frac{1}{2}.$$

To arrive at the equation of the perpendicular from the point P to the line AB , we shall write the equation of an arbitrary line through the point P ; by formula (7) of Art. 56, we get

$$y - 9 = k(x - 4), \quad (*)$$

where k is the slope (unspecified as yet). The required line must be perpendicular to AB ; consequently, its slope must satisfy the condition for perpendicularity with respect to the line AB . The slope of AB being equal to $\frac{1}{2}$, it follows from formula (2) that $k = -2$. Inserting this value of k in equation (*), we obtain

$$y - 9 = -2(x - 4) \text{ or } y = -2x + 17.$$

Solving simultaneously the equations

$$y = \frac{1}{2}x - \frac{1}{2},$$

$$y = -2x + 17,$$

we find the coordinates of the projection sought:

$$x = 7, y = 3.$$

§ 19. The Straight Line As the Curve of the First Order. The General Equation of the Straight Line

60. We shall now establish the following fundamental

Theorem 9. *In cartesian coordinates, every straight line is represented by an equation of the first degree and, conversely, every equation of the first degree represents a straight line.*

Proof. We shall begin by proving the first part of the theorem. Let there be given an arbitrary straight line. If this line is not perpendicular to the axis Ox , then, according to Art. 53, it is represented by an equation of the form $y = kx + b$, that is, by an equation of the first degree.

If the line is perpendicular to the axis Ox , then all its points have the same abscissa, equal to the x -intercept of the line (that is, to the value of the segment cut off by the line on the axis Ox ; see Fig. 40); denoting this x -intercept by the letter a , we thus obtain the equation of the line in the form $x = a$ which, again, is an equation of the first degree. Hence, every straight line is represented by an equation of the first degree in cartesian coordinates; the first part of the theorem is thus proved.

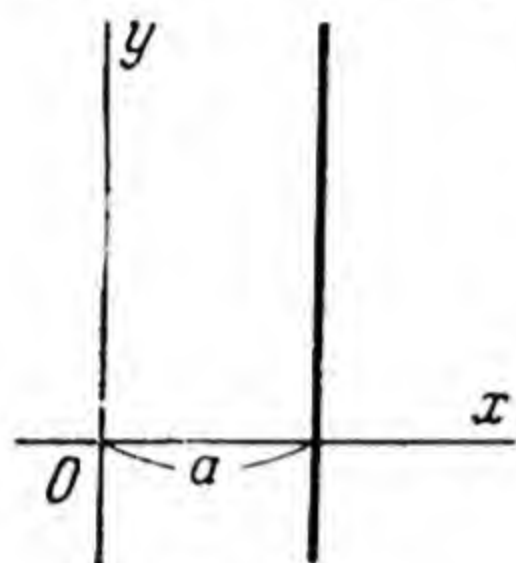


Fig. 40.

We now proceed to prove the converse statement. Let there be given an equation of the first degree,

$$Ax + By + C = 0, \quad (1)$$

where A, B, C may have any values whatsoever. If $B \neq 0$, the given equation may be written in the form

$$y = -\frac{A}{B}x - \frac{C}{B}.$$

Replacing $-\frac{A}{B}$ by k and $-\frac{C}{B}$ by b , we obtain $y = kx + b$; by Art. 53, an equation of this form represents a straight line with slope k and y -intercept b .

If $B = 0$, then $A \neq 0$, and equation (1) may be put in the form

$$x = -\frac{C}{A}.$$

Denoting $-\frac{C}{A}$ by a , we get $x = a$, that is, the equation of a line perpendicular to the axis Ox . Thus, every equation of the first degree represents a straight line. This completes the proof of the theorem.

We know that curves represented by equations of the first degree in cartesian coordinates are called curves of the first order (see Art. 48). Using this term, we can express the above result as follows: *Every straight line is a curve of the first order; every curve of the first order is a straight line.*

61. An equation of the form $Ax + By + C = 0$ is called *the general equation of a straight line* (inasmuch as it is a general

equation of the first degree). By varying the values of A , B , C , it can be made to represent every possible straight line without exception.

§ 20. Incomplete Equations of the First Degree. The Intercept Equation of a Straight Line

62. Let us consider three special cases when the equation of the first degree is incomplete.

(1) $C = 0$; the equation has the form $Ax + By = 0$ and represents a straight line through the origin.

For, the numbers $x = 0$, $y = 0$ satisfy the equation $Ax + By = 0$. Consequently, our straight line contains the origin.

(2) $B = 0$ ($A \neq 0$); the equation has the form $Ax + C = 0$ and represents a line parallel to the axis Oy .

This case has already been discussed in Art. 60, in the process of proving Theorem 9. As shown there, the equation $Ax + C = 0$ is reducible to the form

$$x = a,$$

where $a = -\frac{C}{A}$. An equation of this form represents a straight line perpendicular to the axis Ox because, according to this equation, all points of the line have the same abscissa ($x = a$) and are therefore situated at the same distance from the axis Oy (to the "right" of Oy if the x -intercept a of the line is positive, or to the "left" of Oy if a is negative; see Fig. 40).

In particular, when $a = 0$, the straight line coincides with the axis Oy . Thus, the equation

$$x = 0$$

represents the y -axis.

(3) $A = 0$ ($B \neq 0$); the equation assumes the form $By + C = 0$ and represents a straight line parallel to the axis Ox .

The proof is analogous to that used in the previous case. It will suffice to note that, letting $-\frac{C}{B} = b$, the equation $By + C = 0$ may be written as

$$y = b,$$

where the number b is the "level of location" common to all points of the line (Fig. 41), and also the y -intercept of the line.

In particular, when $b = 0$, the straight line coincides with the axis Ox . Thus, the equation

$$y = 0$$

represents the x -axis.

63. Let us now consider the equation

$$Ax + By + C = 0,$$

in which the coefficients A , B , C are all different from zero. Such an equation can be reduced to a special form, which is found convenient when dealing with some problems of analytic geometry. Transposing the constant term C to the right-hand side of the equation, we obtain

$$Ax + By = -C.$$

Dividing the equation through by $-C$, we then get

$$\frac{Ax}{-C} + \frac{By}{-C} = 1,$$

or

$$\frac{x}{-\frac{C}{A}} + \frac{y}{-\frac{C}{B}} = 1.$$

Introducing the notation

$$a = -\frac{C}{A}, \quad b = -\frac{C}{B},$$

we have

$$\frac{x}{a} + \frac{y}{b} = 1. \quad (1)$$

This is the special form of the equation of a straight line that we wished to obtain.

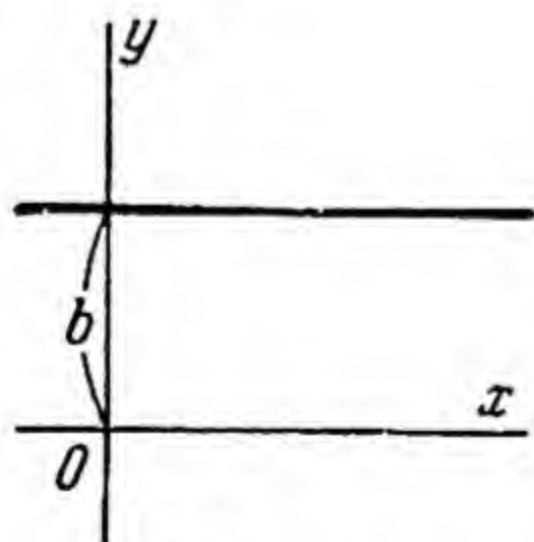


Fig. 41.

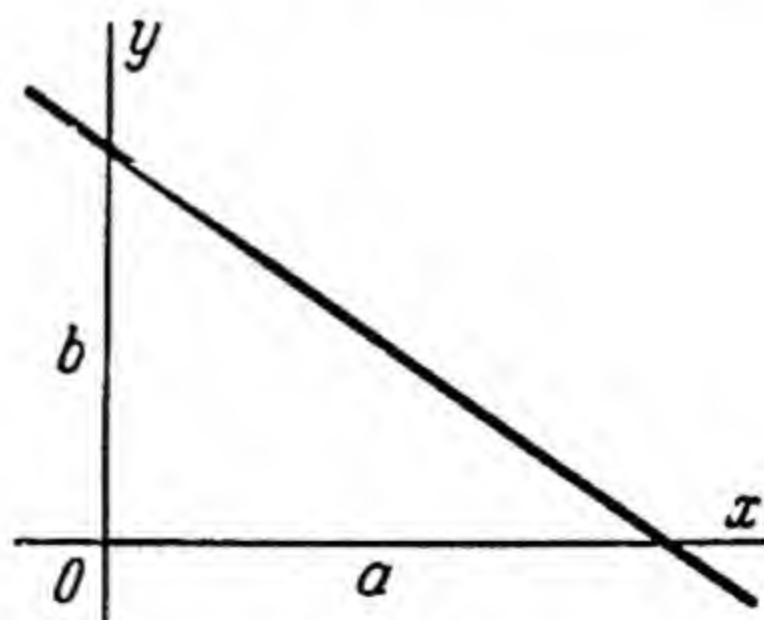


Fig. 42.

It is an important fact that the numbers a and b have a very simple geometric meaning; namely, a and b are the intercepts cut off by the line on the x - and y -axes, respectively (see Fig. 42). To verify this, let us find the points in which our line meets the coordinate axes. The point of intersection of the line and the axis

Ox is determined by solving simultaneously the equations of the line and of the axis Ox :

$$\left. \begin{aligned} \frac{x}{a} + \frac{y}{b} &= 1, \\ y &= 0. \end{aligned} \right\}$$

Hence $x = a$, $y = 0$. Thus, the x -intercept of the line is actually equal to a . In a similar way, the y -intercept of the line is shown to be equal to b .

An equation of the form (1) is called *the intercept equation of a straight line*. This form is, in particular, convenient to use when plotting straight lines on paper.

Example. Given the line

$$3x - 5y + 15 = 0.$$

Write its equation in the intercept form and draw the line.

Solution. The intercept form of the equation of the given line is

$$\frac{x}{-5} + \frac{y}{3} = 1.$$

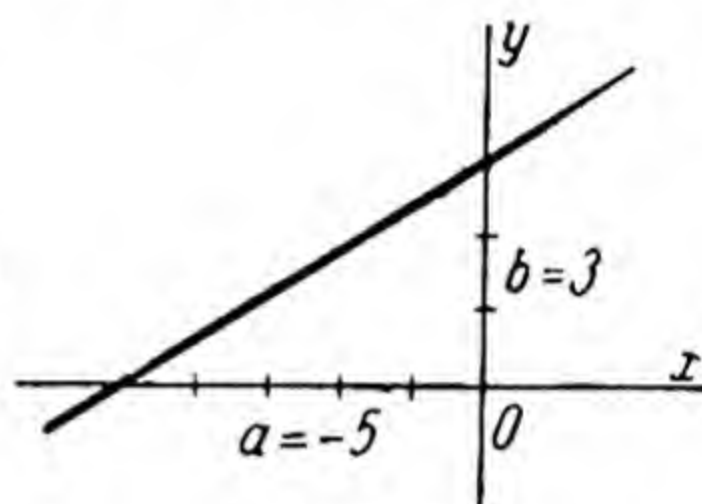


Fig. 43.

To plot the line, mark off the x -intercept $a = -5$ and the y -intercept $b = 3$ on the coordinate axes Ox and Oy , respectively, and join the points so found (Fig. 43).

§ 21. Discussion of a System of Equations Representing Two Straight Lines

64. Let there be given a system of two equations of the first degree:

$$\left. \begin{aligned} A_1x + B_1y + C_1 &= 0, \\ A_2x + B_2y + C_2 &= 0. \end{aligned} \right\} \quad (1)$$

Each of equations (1), taken separately, represents a straight line. Each simultaneous solution of these equations represents a point common to the two straight lines.

We shall analyse the system (1) and give a geometric interpretation to the results of our analysis.

Suppose that $\frac{A_1}{A_2} \neq \frac{B_1}{B_2}$. In this case the determinant of the system is different from zero:

$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \neq 0.$$

This means that the system is consistent and has a single solution *); accordingly, the lines represented by the equations of the system intersect in a single point; hence, *these straight lines are distinct and non-parallel*. The coordinates of the point of intersection are found from equations (1) by means of the formulas

$$x = \frac{\begin{vmatrix} -C_1 & B_1 \\ -C_2 & B_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} A_1 & -C_1 \\ A_2 & -C_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}$$

or

$$x = \frac{\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix}}{\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix}}. \quad (2)$$

Suppose now that $\frac{A_1}{A_2} = \frac{B_1}{B_2}$. Here, again, two cases are possible: either $\frac{A_1}{A_2} = \frac{B_1}{B_2} \neq \frac{C_1}{C_2}$, or $\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$.

Consider the first of these cases. Denoting each of the equal ratios $\frac{A_1}{A_2}$ and $\frac{B_1}{B_2}$ by the letter q , we may write: $A_1 = A_2q$, $B_1 = B_2q$, $C_1 \neq C_2q$. Multiplying the second of equations (1) through by q and subtracting the result from the first equation, we obtain $C_1 - C_2q = 0$. This relation is in contradiction to $C_1 \neq C_2q$. Still, it is a consequence of the system (1); hence, no matter what may be the values of the variables x , y , the equations of the system (1) cannot give correct equalities simultaneously, that is, the system (1) has no simultaneous solutions. In this case, *the straight lines represented by equations (1) have no point in common, that is, are parallel*.

Consider now the other possible case, in which

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}.$$

Equating each of these ratios to q yields $A_1 = A_2q$, $B_1 = B_2q$, $C_1 = C_2q$. Multiplying the left-hand side of the second equation by a certain number q , we therefore obtain the left-hand side of the first equation. It follows that equations (1) are equivalent. Hence, *both equations (1) represent the same straight line*.

* See Art. 2 of the Appendix.

Examples. (1) The lines

$$3x + 4y - 1 = 0,$$

$$2x + 3y - 1 = 0$$

intersect, since $\frac{3}{2} \neq \frac{4}{3}$. The coordinates of the point of intersection are $x = -1, y = +1$.

(2) The lines

$$2x + 3y + 1 = 0,$$

$$4x + 6y + 3 = 0$$

are parallel, because $\frac{2}{4} = \frac{3}{6} \neq \frac{1}{3}$. (The given system is obviously inconsistent, since multiplying the first equation through by 2 and subtracting the result from the second equation gives the contradictory equality $1 = 0$.)

(3) The lines

$$x + y + 1 = 0,$$

$$2x + 2y + 2 = 0$$

coincide, because the given equations are equivalent.

Note. The relation $\frac{A_1}{A_2} = \frac{B_1}{B_2}$ is known as the condition for the parallelism of two straight lines

$$A_1x + B_1y + C_1 = 0 \quad \text{and} \quad A_2x + B_2y + C_2 = 0,$$

although under this condition the lines may be either *parallel* or *coincident*, as we have seen above. Before calling $\frac{A_1}{A_2} = \frac{B_1}{B_2}$ the condition for the parallelism of two straight lines, we must therefore agree to regard the case when the two lines coincide as a special (limiting) case of their parallelism.

65. As an immediate result of the above discussion, we have the following important proposition:

Two equations,

$$A_1x + B_1y + C_1 = 0 \quad \text{and} \quad A_2x + B_2y + C_2 = 0,$$

represent the same straight line if, and only if, their coefficients are proportional, that is,

$$\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}.$$

We shall use this result later.

§ 22. The Normal Equation of a Straight Line. The Problem of Calculating the Distance of a Point from a Straight Line

66. We shall now consider a special form, known as the *normal* form, of the equation of a straight line.

Let there be given an arbitrary straight line. Through the origin, draw the line n (called the *normal*) perpendicular to the given line, and denote by the letter P the point at which the normal cuts the given line (Fig. 44).

We shall regard the direction from the point O to the point P as the positive direction of the normal (if P coincides with O , that is, if the given line passes through the origin, the positive direction of the normal may be chosen at will). Thus, the normal will be regarded as an axis.

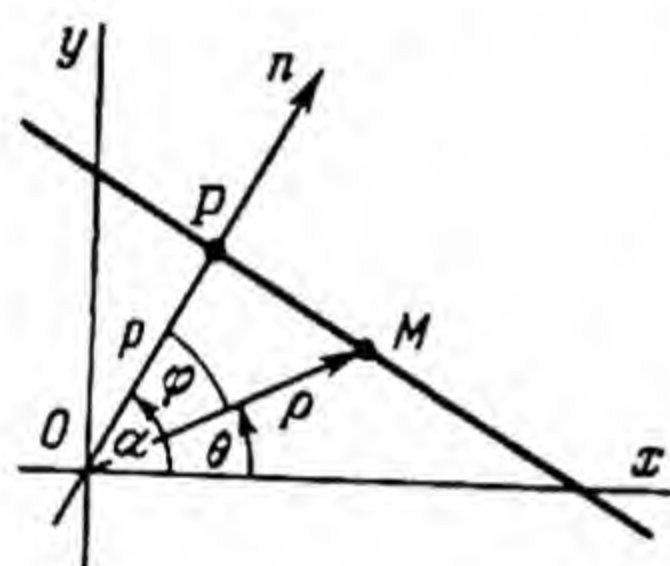


Fig. 44.

Let α be the angle from the axis Ox to the directed normal, and let p denote the length of the segment \overline{OP} .

The angle α will be considered as in trigonometry and called the polar angle of the normal.

We shall now derive the equation of the given line, assuming that the numbers α and p are known. For this purpose, take an arbitrary point M on the line and designate its coordinates as x, y ; obviously, the projection of the segment \overline{OM} on the normal is equal to \overline{OP} and, since the positive direction of the normal agrees with that of the segment \overline{OP} , the value of the segment is represented by a positive number, namely, the number p :

$$\text{proj}_n \overline{OM} = p. \quad (1)$$

Let us find the expression for the projection of the segment \overline{OM} on the normal in terms of the coordinates of the point M . Letting φ denote the angle between the segment \overline{OM} and the normal, and ρ, θ the polar coordinates of the point M , we obtain, by Art. 20,

$$\begin{aligned} \text{proj}_n \overline{OM} &= \rho \cos \varphi = \rho \cos (\alpha - \theta) = \rho (\cos \alpha \cos \theta + \sin \alpha \sin \theta) = \\ &= (\rho \cos \theta) \cos \alpha + (\rho \sin \theta) \sin \alpha = x \cos \alpha + y \sin \alpha. \end{aligned}$$

Thus,

$$\text{proj}_n \overline{OM} = x \cos \alpha + y \sin \alpha. \quad (2)$$

From (1) and (2) it follows that $x \cos \alpha + y \sin \alpha = p$, or

$$x \cos \alpha + y \sin \alpha - p = 0. \quad (3)$$

This is the desired equation of the given line (as we see, it is satisfied by the coordinates x, y of every point M lying on the line; on the other hand, if a point M does not lie on the line, its coordinates do not satisfy equation (3), because then $\text{proj}_n \overline{OM} \neq p$).

The equation of a straight line written in the form (3) is called the normal equation of a straight line; α denotes here the polar angle of the normal, and p is the distance from the origin to the straight line.

67. Let there be given an arbitrary straight line. Construct its normal n (assigning a positive direction to it, as described in the preceding article). Further, let M^* be any point in the plane, and let d denote the distance of M^* from the given line (Fig. 45).

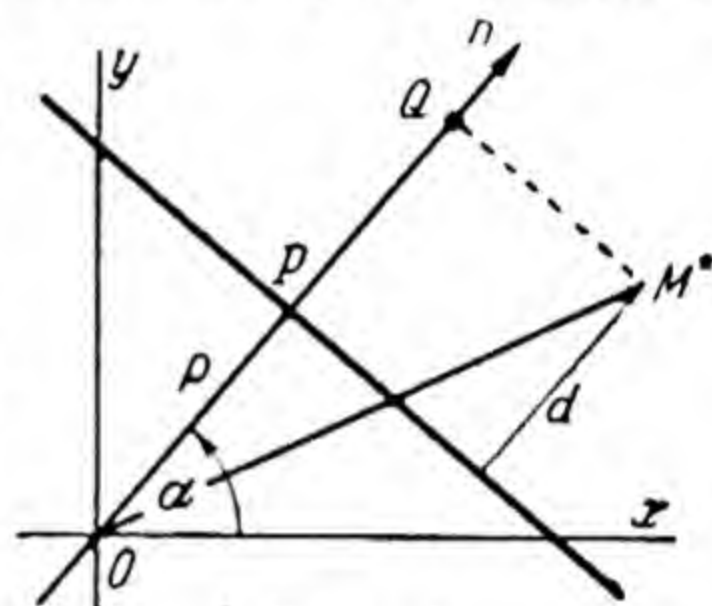


Fig. 45.

We shall agree to define the departure of the point M^* from the given straight line as the number $+d$ if M^* lies on that side of the line towards which the directed normal points, and as the number $-d$ if M^* lies on the other side of the line. We shall denote the departure of a point from a straight

line by the letter δ ; thus, $\delta = \pm d$, and it will be helpful to note that $\delta = +d$ when the point M^* and the origin are on opposite sides of the line, and $\delta = -d$ when M^* and the origin are on the same side of the line. (For points lying on the line, $\delta = 0$.)

The problem of calculating the departure of a point from a straight line is one of the standard problems of analytic geometry. This problem is solved by means of the following

Theorem 10. If a point M^* has coordinates (x^*, y^*) and a straight line is represented by the normal equation

$$x \cos \alpha + y \sin \alpha - p = 0,$$

the departure of the point M^* from the straight line is given by the formula

$$\delta = x^* \cos \alpha + y^* \sin \alpha - p. \quad (4)$$

Proof. Project the point M^* on the normal and denote the projection by Q (Fig. 45). We have

$$\delta = PQ = OQ - OP,$$

where PQ , OQ and OP are the values of the directed segments \overline{PQ} , \overline{OQ} and \overline{OP} on the normal. But $OQ = \text{proj}_n \overline{OM}^*$, $OP = p$; hence

$$\delta = \text{proj}_n \overline{OM}^* - p. \quad (5)$$

Applying formula (2) of Art. 66 to our point M^* , we get

$$\text{proj}_n \overline{OM}^* = x^* \cos \alpha + y^* \sin \alpha. \quad (6)$$

From (5) and (6), we have

$$\delta = x^* \cos \alpha + y^* \sin \alpha - p.$$

The theorem is thus proved.

Note now that $x^* \cos \alpha + y^* \sin \alpha - p$ is nothing more than the left-hand member of the normal equation of the given straight line, with the current coordinates replaced by the coordinates of the point M^* . We hence obtain the following rule:

To find the departure of a point M^ from a straight line, the coordinates of the point M^* must be substituted for the current coordinates in the left-hand member of the normal equation of the straight line. The resulting number will be the departure required.*

Note. The distance of a point from a straight line is equal to the modulus (the absolute value) of the departure of that point: $d = |\delta|$. Consequently, to find the distance of a point from a straight line, it is sufficient to calculate the departure by the rule just given and to take the modulus of this departure.

68. As we have seen, the problem of calculating the departure of a point from a straight line is readily solved if the straight line is represented by its *normal* equation. It will now be shown how to reduce the general equation of a straight line to the normal form. Let

$$Ax + By + C = 0 \quad (7)$$

be the general equation of some straight line, and let

$$x \cos \alpha + y \sin \alpha - p = 0 \quad (3)$$

be its normal equation.

Since equations (7) and (3) represent the same line, the coefficients of these equations are proportional, according to Art. 65. This means that, on multiplying equation (7) throughout by a certain factor μ , we shall obtain the equation

$$\mu Ax + \mu By + \mu C = 0,$$

which will be identical with equation (3); that is, we shall have

$$\mu A = \cos \alpha, \quad \mu B = \sin \alpha, \quad \mu C = -p. \quad (8)$$

To find the factor μ , square and add the first two of these relations; this gives

$$\mu^2(A^2 + B^2) = \cos^2 \alpha + \sin^2 \alpha = 1.$$

Hence,

$$\mu = \frac{\pm 1}{\sqrt{A^2 + B^2}}. \quad (9)$$

The number μ , multiplication by which reduces the general equation of a line to the normal form, is called the *normalising factor* of that equation. Formula (9) determines the normalising factor incompletely, since its sign remains undetermined.

To determine the sign of the normalising factor, let us use the third of relations (8). According to this relation, μC is a negative number. Hence, *the normalising factor is opposite in sign to the constant term of the equation normalised.*

Note. If $C=0$, the sign of the normalising factor may be chosen at pleasure.

Example. Given the line $3x - 4y + 10 = 0$ and the point $M(4, 3)$. Find the departure of the point M from the given line.

Solution. To apply the rule stated in Art. 67, we must first reduce the given equation to its normal form.

To this end, we find the normalising factor

$$\mu = \frac{-1}{\sqrt{3^2 + 4^2}} = -\frac{1}{5}.$$

Multiplying our equation by μ , we obtain the required normal equation

$$-\frac{1}{5}(3x - 4y + 10) = 0.$$

Substituting the coordinates of the point M in the left member of this equation, we obtain

$$d = -\frac{1}{5}(3 \cdot 4 - 4 \cdot 3 + 10) = -2.$$

Thus, the point M has a negative departure from the given line and is at the distance $d = 2$ from it.

§ 23. The Equation of a Pencil of Lines

69. The collection of all those lines in the plane which pass through a point $S(x_0, y_0)$ is called *the pencil* of lines with vertex S . In analytic geometry it is often necessary to find, from the known equations of two lines of a pencil, the equation of a third line of the same pencil, provided that the direction of this third line has been specified in some way. Problems of this type can be solved by using, for instance, equation (7) of Art. 56:

$y - y_1 = k(x - x_1)$, with the coordinates x_0, y_0 of the vertex of the pencil taken as x_1, y_1 (the slope k is determined here according to the manner in which the direction of the required line has been specified). When employing this method, however, the coordinates x_0, y_0 of the vertex must first be computed.

The following proposition permits us to solve such problems without computing the coordinates x_0, y_0 .

Let $A_1x + B_1y + C_1 = 0$, $A_2x + B_2y + C_2 = 0$ be the equations of two straight lines intersecting in the point S , and let α, β be any numbers which are not both simultaneously equal to zero; then

$$\alpha(A_1x + B_1y + C_1) + \beta(A_2x + B_2y + C_2) = 0 \quad (1)$$

is the equation of a line through the point S .

Proof. Let us, first of all, establish that relation (1) is actually an equation (rather than an identity). For this purpose, we put it in the form

$$(\alpha A_1 + \beta A_2)x + (\alpha B_1 + \beta B_2)y + (\alpha C_1 + \beta C_2) = 0 \quad (2)$$

and proceed to show that the quantities $\alpha A_1 + \beta A_2$ and $\alpha B_1 + \beta B_2$ cannot both be zero. Suppose the converse is true, that is, $\alpha A_1 + \beta A_2 = 0$ and $\alpha B_1 + \beta B_2 = 0$; but then $\frac{A_1}{A_2} = -\frac{\beta}{\alpha}$ and $\frac{B_1}{B_2} = -\frac{\beta}{\alpha}$. Since the numbers α and β are not both zero, the ratio $\frac{\beta}{\alpha}$ cannot be indeterminate; the last two relations therefore yield the proportion $\frac{A_1}{A_2} = \frac{B_1}{B_2}$. But the coefficients A_1, B_1 cannot be proportional to the coefficients A_2, B_2 , since the given lines intersect (see Art. 64). Hence, our supposition has to be rejected. Thus, $\alpha A_1 + \beta A_2$ and $\alpha B_1 + \beta B_2$ cannot vanish simultaneously, which means that (2) is an equation (in the variables x and y). Further, it is immediately evident that (2) is a *first-degree equation* and, hence, represents a straight line. It remains to prove that this line passes through the point S . Let x_0, y_0 be the coordinates of S . Since each of the given lines passes through the point S , it follows that $A_1x_0 + B_1y_0 + C_1 = 0$ and $A_2x_0 + B_2y_0 + C_2 = 0$, whence

$$\alpha(A_1x_0 + B_1y_0 + C_1) + \beta(A_2x_0 + B_2y_0 + C_2) = 0.$$

We see that the coordinates of the point S satisfy equation (1); consequently, the line represented by (1) passes through S , and so the proof is complete.

Thus, an equation of the form (1) represents (for any values of α, β , not both zero) a straight line of the pencil with vertex S .

Let us now prove that *it is always possible to choose the numbers α, β so as to make equation (1) represent any (previously assigned) line of the pencil with vertex S* . Since each line of the pencil with vertex S is determined by specifying the point S and one more point on the line, it follows that, to prove the assertion just made, we have merely to show that the numbers α, β in (1) can always be chosen so as to make the line represented by (1) pass through any preassigned point $M^*(x^*, y^*)$.

But this is evident; for, the line represented by (1) will pass through a point M^* if the coordinates of M^* satisfy this equation, that is, if

$$\alpha(A_1x^* + B_1y^* + C_1) + \beta(A_2x^* + B_2y^* + C_2) = 0. \quad (3)$$

We assume that the point M^* does not coincide with the point S (this being the only case we are concerned with). Then at least one of the numbers

$$A_1x^* + B_1y^* + C_1, \quad A_2x^* + B_2y^* + C_2$$

is different from zero, so that (3) is an equation, rather than an identity; namely, (3) is an equation of the first degree in two unknowns, α and β . To find the unknowns α, β , one of them is assigned an arbitrarily chosen value, and then the value of the other is computed from the equation; for instance, if $A_2x^* + B_2y^* + C_2 \neq 0$, then α may be assigned any value (other than zero) and the corresponding value of β determined from the relation

$$\beta = -\frac{A_1x^* + B_1y^* + C_1}{A_2x^* + B_2y^* + C_2} \alpha.$$

Thus, an equation of the form (1) can be made to represent a straight line passing through any preassigned point of the plane and, hence, to represent any straight line of the pencil with vertex S . An equation of the form (1) is therefore called *the equation of a pencil of lines (with vertex S)*.

If $\alpha \neq 0$, then, letting $\frac{\beta}{\alpha} = \lambda$, we obtain from (1):

$$A_1x + B_1y + C_1 + \lambda(A_2x + B_2y + C_2) = 0. \quad (4)$$

In problem-solving practice, this form of the equation of a pencil of lines is used more frequently than the form (1). However, it is important to note that, since the case $\alpha = 0$ is eliminated when reducing (1) to (4), an equation of the form (4) cannot represent the line $A_2x + B_2y + C_2 = 0$; that is, an equation of the form (4) can be made, by varying the value of λ , to represent every

straight line of the pencil except one (the second of the two given lines).

Example. Given the two lines $2x + 3y - 5 = 0$, $7x + 15y + 1 = 0$, which intersect in the point S . Find the equation of the line through S and perpendicular to the line $12x - 5y - 1 = 0$.

Solution. Let us first verify the data: the given lines do intersect, because $\frac{2}{7} \neq \frac{3}{15}$. Now, we write the equation of the pencil of lines with vertex S :

$$2x + 3y - 5 + \lambda(7x + 15y + 1) = 0. \quad (5)$$

To single out the required line from this pencil of lines, let us compute λ according to the condition that the line in question must be perpendicular to the line $12x - 5y - 1 = 0$. Rewriting equation (5) in the form

$$(2 + 7\lambda)x + (3 + 15\lambda)y + (-5 + \lambda) = 0, \quad (6)$$

we find the slope of the required line:

$$k = -\frac{2 + 7\lambda}{3 + 15\lambda}.$$

Now, the given line has the slope

$$k_1 = \frac{12}{5}.$$

By the perpendicularity condition, $k = -\frac{1}{k_1}$, that is,

$$-\frac{2 + 7\lambda}{3 + 15\lambda} = -\frac{5}{12}.$$

Hence $\lambda = -1$. Substituting $\lambda = -1$ in (6), we get

$$-5x - 12y - 6 = 0,$$

or

$$5x + 12y + 6 = 0.$$

The problem is solved.

GEOMETRIC PROPERTIES OF CURVES OF THE SECOND ORDER

The present chapter deals with the three types of second-order curves: *the ellipse, hyperbola and parabola*. The main object of the chapter is to acquaint the reader with the more important geometric properties of these curves.

§ 24. The Ellipse. Definition of the Ellipse and Derivation of Its Canonical Equation

70. *An ellipse is the locus of points the sum of whose distances from two fixed points (called the foci) of the plane is a constant; this constant is required to be greater than the distance between the foci.* It is customary to designate the foci of an ellipse as F_1 and F_2 .

Note. It is obvious that the sum of the distances of an arbitrary point M from two fixed points F_1 and F_2 cannot be less than the distance between the points F_1 and F_2 . This sum is equal to the distance between F_1 and F_2 if, and only if, the point M lies on the line segment F_1F_2 . Consequently, the locus of points, the sum of whose distances from the two fixed points F_1 and F_2 is a constant *equal* to the distance between F_1 and F_2 , is the segment F_1F_2 itself; this case has been excluded by the restriction made at the end of the above definition.

71. Let M be an arbitrary point of an ellipse with foci F_1, F_2 . The segments F_1M and F_2M (as well as their lengths) are called *the focal radii* of the point M . The constant sum of the focal radii of a point on an ellipse is generally denoted by $2a$. Thus, for any point M of an ellipse,

$$F_1M + F_2M = 2a. \quad (1)$$

The distance F_1F_2 between the foci is denoted by $2c$. From

$$F_1M + F_2M > F_1F_2,$$

we have

$$2a > 2c, \quad \text{that is,} \quad a > c. \quad (2)$$

The following method of constructing an ellipse by means of a piece of thread is directly based on the definition of the ellipse. Fasten at points F_1 and F_2 the ends of an inextensible thread of length $2a$ and stretch the thread taut with the point of a pencil. Move the pencil point and it will describe an ellipse with F_1, F_2 as the foci and $2a$ as the sum of focal radii. On completing the actual construction, it will be clearly seen that the ellipse is a convex closed curve (an oval) symmetric with respect to the line F_1F_2

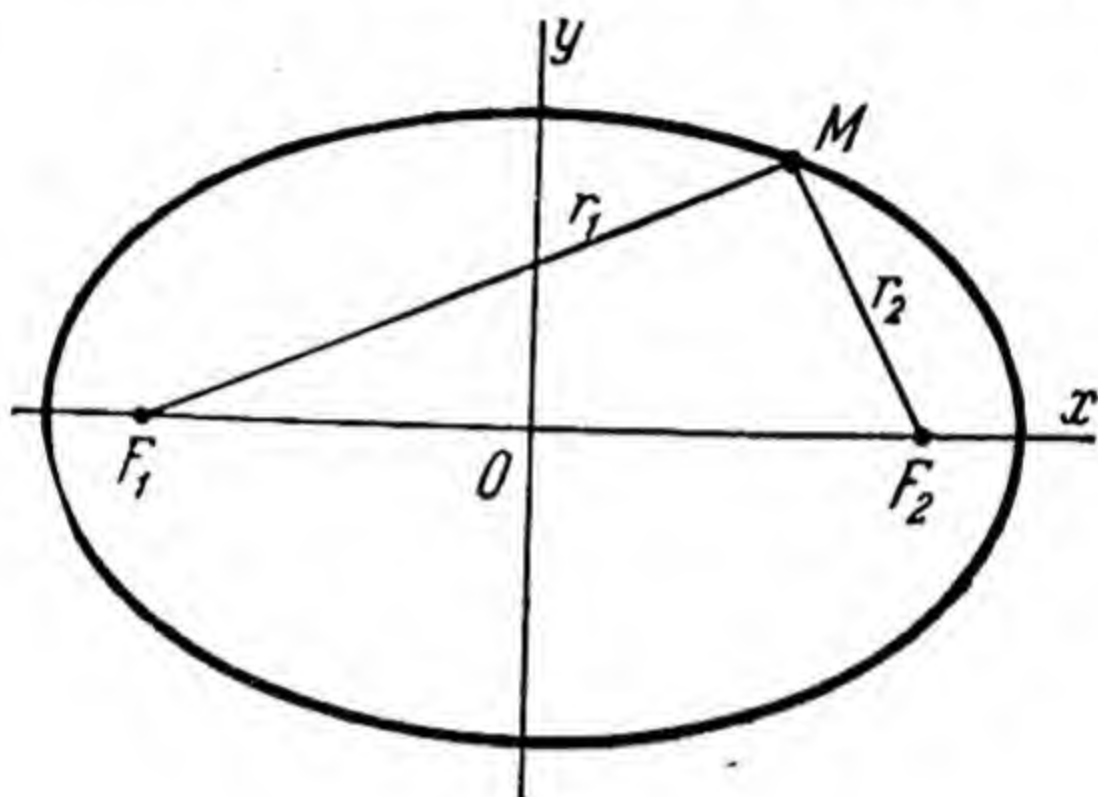


Fig. 46.

and also with respect to the perpendicular bisector of the segment F_1F_2 (Fig. 46). A little later, we shall establish the shape of the ellipse analytically by discussing its equation derived in the next article.

72. Let there be given an ellipse with foci F_1, F_2 (we assume that a and c are also given). Let us attach to the plane a rectangular cartesian coordinate system, whose axes are specially chosen with respect to the ellipse; namely, let the line F_1F_2 be taken as the x -axis, the direction from F_1 to F_2 adopted as positive, and the origin placed at the midpoint of the segment F_1F_2 (Fig. 46). We proceed to derive the equation of the ellipse referred to the chosen coordinate system.

Take an arbitrary point M in the plane. Designate its coordinates as x and y , and the distances of the point M from the foci as r_1 and r_2 ($r_1 = F_1M$, $r_2 = F_2M$). The point M will lie on the given ellipse if, and only if,

$$r_1 + r_2 = 2a. \quad (3)$$

In order to obtain the desired equation, it is necessary to express the variables r_1 and r_2 in terms of the coordinates x, y and to substitute these expressions in (3).

Note that the coordinates of the foci F_1 and F_2 are $(-c, 0)$ and $(+c, 0)$, respectively, since $F_1F_2 = 2c$ and since the foci are symmetrically situated on the axis Ox with respect to the origin; bearing this in mind and using formula (2) of Art. 18, we find

$$r_1 = \sqrt{(x+c)^2 + y^2}, \quad r_2 = \sqrt{(x-c)^2 + y^2}. \quad (4)$$

Substituting these expressions for r_1 and r_2 in (3), we obtain

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a. \quad (5)$$

The coordinates of a point $M(x, y)$ satisfy equation (5) if, and only if, the point M lies on the given ellipse; consequently, (5) is the equation of this ellipse in the chosen coordinate system. The purpose of the remaining operations is to arrive at a simpler form of the equation of the ellipse.

Transposing the second radical to the right side of equation (5) and squaring both members, we obtain

$$(x+c)^2 + y^2 = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2, \quad (6)$$

or

$$a\sqrt{(x-c)^2 + y^2} = a^2 - cx. \quad (7)$$

Squaring both members of the last relation yields

$$a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2 = a^4 - 2a^2cx + c^2x^2, \quad (8)$$

whence

$$(a^2 - c^2)x^2 + a^2y^2 = a^2(a^2 - c^2). \quad (9)$$

We shall here introduce a new quantity

$$b = \sqrt{a^2 - c^2}; \quad (10)$$

the geometric meaning of the quantity b will be clarified a little later; we shall only note now that b is a real quantity (since $a > c$ and, consequently, $a^2 - c^2 > 0$). From (10), we have

$$b^2 = a^2 - c^2; \quad (11)$$

hence, equation (9) may be written as

$$b^2x^2 + a^2y^2 = a^2b^2,$$

or

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (12)$$

Let us prove that *equation (12) is the equation of the given ellipse*. This is not a self-evident fact since we have twice cleared radicals in the process of reducing (5) to (12); it is obvious only that (12) is a consequence of (5). We must prove that (5) is, in

its turn, a consequence of (12), i. e., that these equations are equivalent.

Let x and y be any two numbers for which equation (12) is valid. Reversing the steps in the above derivation, we return from (12) to (9), and then to (8), which will now be written in the form

$$a^2[(x-c)^2 + y^2] = (a^2 - cx)^2.$$

Extracting the square root of both sides of the equation, we get

$$a\sqrt{(x-c)^2 + y^2} = \pm(a^2 - cx). \quad (13)$$

Note that, in virtue of (12), $|x| \leq a$. From this and from the fact that $c < a$, it follows that $|cx| < a^2$; hence $a^2 - cx$ is a positive number. Accordingly, the right side of (13) must be taken with the plus sign. This brings us back to (7), after which we obtain (6); we shall write this last equation in the form

$$(x+c)^2 + y^2 = [2a - \sqrt{(x-c)^2 + y^2}]^2.$$

Hence

$$\sqrt{(x+c)^2 + y^2} = \pm(2a - \sqrt{(x-c)^2 + y^2}). \quad (14)$$

Let us discuss the value of

$$(x-c)^2 + y^2 = x^2 - 2cx + c^2 + y^2. \quad (15)$$

From equation (12), $x^2 \leq a^2$. Furthermore, $|cx| < a^2$, and so the absolute value of $-2cx$ is less than that of $2a^2$. Finally, we can also deduce from (12) that $y^2 \leq b^2$; that is, $y^2 \leq a^2 - c^2$, or $c^2 + y^2 \leq a^2$. Accordingly, the right-hand member of (15) sums up to less than $4a^2$, so that the square root of that member is less than $2a$. Therefore, the quantity enclosed in the parentheses on the right side of (14) is positive, and hence the right side of (14) must be taken with the plus sign. Thus, we obtain

$$\sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y^2},$$

whence equation (5) immediately follows.

Thus, not only equation (12) is derivable from equation (5) but, conversely, (5) can be derived from (12). These two equations are therefore equivalent, which proves that equation (12) is the equation of the given ellipse.

Equation (12) is called the *canonical* equation of the ellipse.

73. The equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

which represents the ellipse in a certain system of rectangular cartesian coordinates, is an equation of the second degree; accordingly, *the ellipse is a curve of the second order.*

§ 25. Discussion of the Shape of the Ellipse

74. The description of the shape of the ellipse given above (Art. 71) was based on visual appraisal. Let us now investigate the shape of the ellipse by analysing its canonical equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (1)$$

We shall begin by emphasising the following algebraic feature of equation (1): *it contains only terms with even powers of the current coordinates.*

To this algebraic property of equation (1) there corresponds an important geometric property of the curve represented by the equation; namely, *the ellipse represented by equation (1) is symmetrical with respect to both the axis Ox and the axis Oy .*

For, if $M(x, y)$ is a point on this ellipse, that is, if the numbers x, y satisfy equation (1), then the numbers $x, -y$ also satisfy equation (1); consequently, the point $M'(x, -y)$ also lies on the ellipse. But the points $M(x, y)$ and $M'(x, -y)$ are symmetrical with respect to the axis Ox . Thus, all points of the ellipse form pairs symmetrical with respect to the axis Ox . In other words, if we fold the drawing along the axis Ox , the upper part of the ellipse will be brought into coincidence with its lower part, which means that the ellipse is symmetric with respect to the axis Ox .

The symmetry of the ellipse with respect to the axis Oy is proved in a completely analogous manner (the proof being based on the fact that, if equation (1) is satisfied by the numbers x, y , it is also satisfied by the numbers $-x, y$).

To investigate the shape of the ellipse, let us express the quantity y as a function of x , by solving (1) for y :

$$y = \pm \sqrt{b^2 \left(1 - \frac{x^2}{a^2}\right)},$$

or

$$y = \pm \frac{b}{a} \sqrt{a^2 - x^2}. \quad (2)$$

Since the ellipse is symmetrical with respect to both coordinate axes, it will be sufficient to consider only its portion contained in the first quadrant.

This portion of the ellipse lies in the upper half-plane and is therefore associated with the plus sign on the right side of (2);

at the same time, since the first-quadrant portion of the ellipse also lies in the right half-plane, $x \geq 0$ for all points of this portion. Thus, our task is to draw the graph of the function

$$y = + \frac{b}{a} \sqrt{a^2 - x^2}, \quad (3)$$

where $x \geq 0$.

Initially, let $x = 0$; for this value of x , $y = b$. The point $B(0, b)$ is the extreme left point of our graph. Now let x increase from zero. It is obvious that the radicand in (3) will decrease

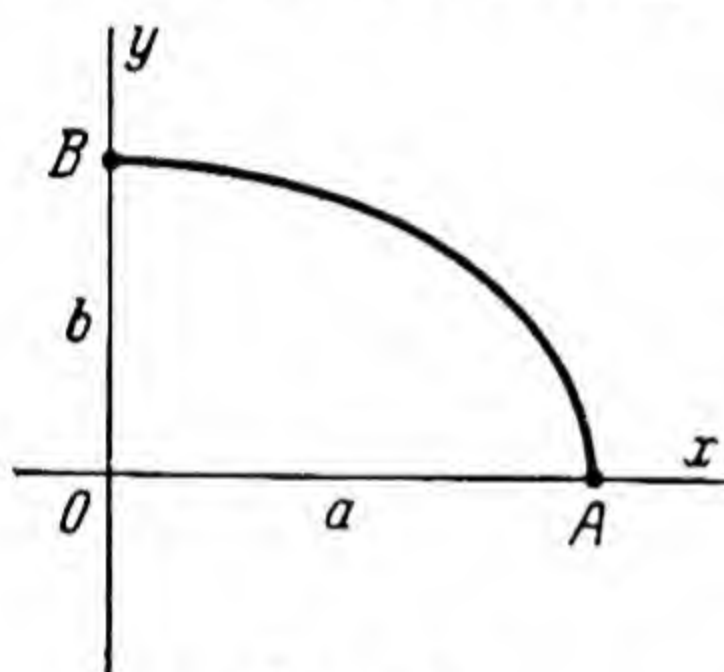


Fig. 47.

as x increases; consequently, y will also decrease in value. Thus, the variable point $M(x, y)$, tracing the graph under consideration, will move to the right and downwards (Fig. 47). When x becomes equal to a , we shall have $y = 0$, and the point $M(x, y)$ will coincide with the point $A(a, 0)$ on the axis Ox . If x increases further, that is, if $x > a$, the radicand in (3) will be negative and y will, therefore, become imaginary. It

follows that the point A is the extreme right point of the graph. Thus, the first-quadrant portion of the ellipse is the arc BA shown in Fig. 47.

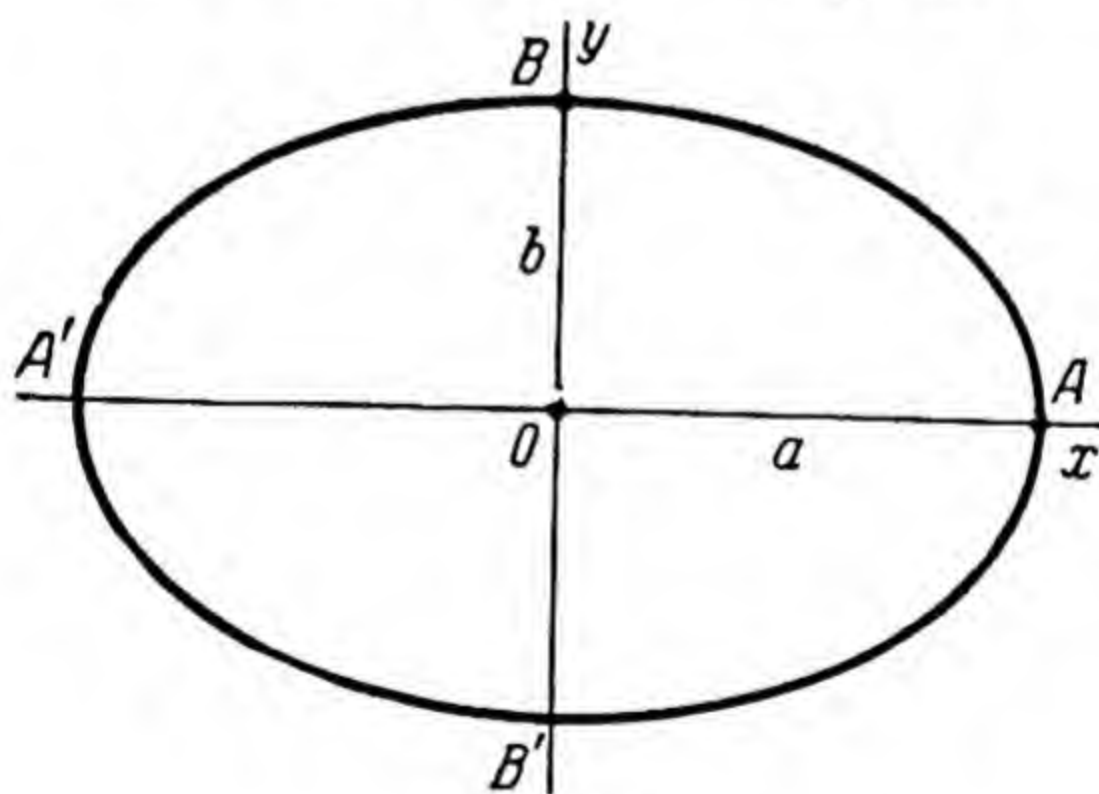


Fig. 48.

By reflecting the arc BA in the coordinate axes, we obtain the entire ellipse; it has the shape of a convex oval with two mutually perpendicular axes of symmetry (Fig. 48).

Usually the axes of symmetry of an ellipse are referred to simply as *the axes* of the ellipse; the intersection of the axes is

called *the centre* of the ellipse. The points where the ellipse cuts its axes are called its *vertices*. In Fig. 48, the vertices of the ellipse are the points A , A' , B , and B' . Note that the segments $AA' = 2a$ and $BB' = 2b$ are also commonly referred to as the axes of the ellipse. If the ellipse is situated relative to the coordinate axes as described in Art. 72, that is, if its foci are on the axis Ox , then $b = \sqrt{a^2 - c^2}$ and, consequently, $a > b$.

In this case, the segment $OA = a$ is called *the semi-major axis*, and the segment $OB = b$, *the semi-minor axis* of the ellipse. But, of course, an ellipse represented by an equation of the form (1) can be placed so that its foci will lie on the axis Oy ; then $b > a$ and the segment $OB = b$ will be the semi-major axis of the ellipse. In either case, however, the length of the segment OA on the x -axis is denoted by a , and the length of the segment OB on the y -axis, by b .

Note. In Fig. 47, the first-quadrant portion of the ellipse is presented as the arc BA , which is convex "upwards" in all points; moreover, in the points B and A the direction of the arc is shown to be perpendicular to the axes Oy and Ox , respectively (so that the full ellipse has no cusps at its vertices). But it remains to be proved that the arc BA actually possesses such properties. However, the proof will be omitted here, since the most convenient methods for graph analysis of such kind are those furnished by the calculus.

75. In the special case where $b = a$, the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

assumes the form

$$x^2 + y^2 = a^2,$$

which is the equation of a circle of radius a (with centre at the origin). Accordingly, the circle is regarded as a special case of the ellipse.

§ 26. The Eccentricity of the Ellipse

76. *The eccentricity of an ellipse is defined as the ratio of the distance between the foci of that ellipse and the length of its major axis; denoting the eccentricity by the letter e , we have*

$$e = \frac{c}{a}.$$

Since $c < a$, it follows that $e < 1$, that is, *the eccentricity of every ellipse is less than unity.*

Note that $c^2 = a^2 - b^2$; therefore

$$\epsilon^2 = \frac{c^2}{a^2} = \frac{a^2 - b^2}{a^2} = 1 - \left(\frac{b}{a}\right)^2;$$

hence

$$\epsilon = \sqrt{1 - \left(\frac{b}{a}\right)^2} \quad \text{and} \quad \frac{b}{a} = \sqrt{1 - \epsilon^2}.$$

Accordingly, the eccentricity is determined by the ratio of the axes of the ellipse and, conversely, the ratio of the axes is determined by the eccentricity. Thus, *the eccentricity characterises the shape of an ellipse*. As the eccentricity increases towards unity, $1 - \epsilon^2$ decreases and, consequently, the ratio $\frac{b}{a}$ diminishes; this means that *the greater the eccentricity, the more elongated is the ellipse*. In the case of a circle, $b = a$ and $\epsilon = 0$.

§ 27. Rational Expressions for Focal Radii of the Ellipse

77. Consider an arbitrary point $M(x, y)$ on a given ellipse. Let r_1 and r_2 be the focal radii of this point; then

$$r_1 = \sqrt{(x+c)^2 + y^2}, \quad r_2 = \sqrt{(x-c)^2 + y^2}. \quad (1)$$

However, the focal radii can also be represented by formulas free of irrational terms. In fact, from equation (7) of Art. 72, we have

$$\sqrt{(x-c)^2 + y^2} = a - \frac{c}{a}x.$$

Letting here $\frac{c}{a} = \epsilon$ and using the second of formulas (1), we obtain

$$r_2 = a - \epsilon x.$$

By the definition of the ellipse,

$$r_1 + r_2 = 2a;$$

substituting $a - \epsilon x$ for r_2 gives

$$r_1 = a + \epsilon x.$$

Thus, we have the formulas

$$\left. \begin{aligned} r_1 &= a + \epsilon x, \\ r_2 &= a - \epsilon x, \end{aligned} \right\} \quad (2)$$

which will find an important application in § 34.

§ 28. Point-by-point Construction of the Ellipse. The Parametric Equations of the Ellipse

78. Let there be given an ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (1)$$

Describe two circles of radii a and b (assuming that $a > b$) about the centre of the ellipse; draw an arbitrary ray from the centre of the ellipse and denote by t the polar angle of this ray (Fig. 49). The ray will cut the larger and the smaller circle in

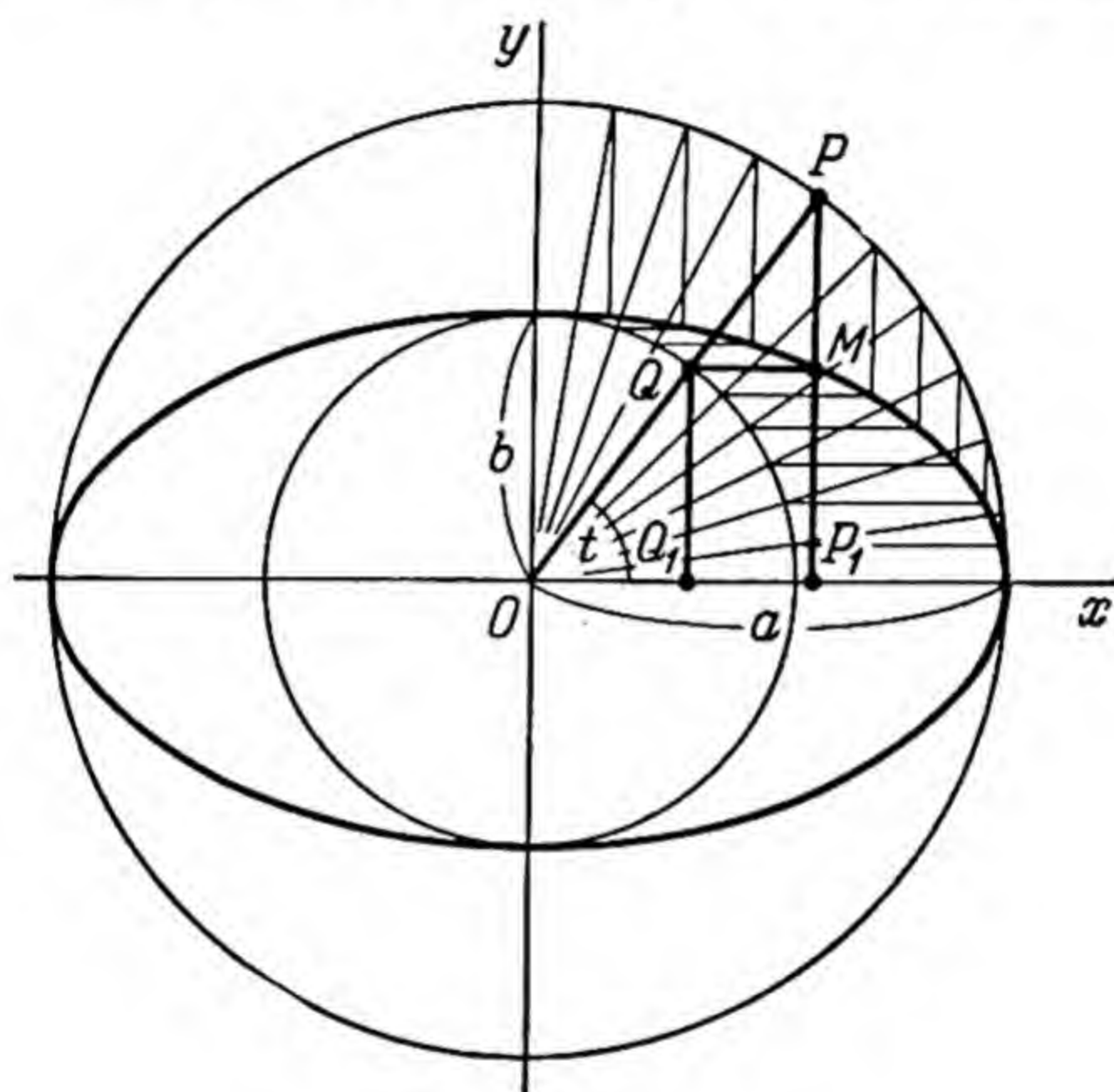


Fig. 49.

points P and Q , respectively. Next, draw a line through the point P parallel to the axis Oy , and another line through the point Q parallel to the axis Ox ; let M be the point of intersection of these lines, and let P_1 and Q_1 be the projections of P and Q on the x -axis.

Let us express the coordinates of the point M in terms of t . From Fig. 49, it is apparent that

$$\begin{aligned} x &= OP_1 = OP \cdot \cos t = a \cos t, \\ y &= P_1M = Q_1Q = OQ \cdot \sin t = b \sin t. \end{aligned}$$

Thus,

$$\left. \begin{aligned} x &= a \cos t, \\ y &= b \sin t. \end{aligned} \right\} \quad (2)$$

Substituting these coordinates in (1), we can see that they satisfy the equation for any value of t . Hence, the point M is on the ellipse. Thus, we have shown how to plot a point of the ellipse. By drawing a number of rays and repeating the construction for each of these rays, we can plot any desired number of points of the ellipse. This method is often used in drawing practice since, by joining the plotted points with the aid of a curved ruler, one can obtain a sketch of the ellipse, fully satisfactory from the practical viewpoint.

79. Equations (2) express the coordinates of an arbitrary point of the ellipse as functions of the variable parameter t ; hence, *equations (2) are the parametric equations of the ellipse* (see § 14).

§ 29. The Ellipse as the Projection of a Circle on a Plane. The Ellipse as the Section of a Circular Cylinder by a Plane

80. We shall now prove that *the projection of a circle on an arbitrary plane is an ellipse*.

Let a circle k lying in a plane β be projected upon a plane α . Denote by k' the locus of the projections of all points of the circle k ; it is to be shown that k' is an ellipse. For convenience let the plane α pass through the centre of the circle k (Fig. 50). Introduce a rectangular cartesian coordinate system in the plane α , taking the line of intersection of the planes α and β as the axis Ox , and the centre of the circle k as the origin. Let a denote the radius of the circle k , and φ the acute angle between the planes α and β . Let P be an arbitrary point of the circle k , M its projection on the plane α , Q its projection on the axis Ox , and t the angle which the segment OP makes with the axis Ox . Now, express the coordinates of the point M in terms of t . From Fig. 50, it is readily seen that

$$\begin{aligned} x &= OQ = OP \cdot \cos t = a \cos t, \\ y &= QM = QP \cdot \cos \varphi = OP \cdot \sin t \cos \varphi = a \cos \varphi \sin t. \end{aligned}$$

Denoting the constant $a \cos \varphi$ by the letter b , we obtain

$$\begin{aligned} x &= a \cos t, \\ y &= b \sin t. \end{aligned}$$

These equations are identical with the parametric equations of the ellipse given in Art. 78; hence, the curve k' is an ellipse (with semi-major axis a and semi-minor axis $b = a \cos \varphi$).

81. Likewise, it can easily be shown that *every section of a circular cylinder by a plane not parallel to the axis of the cylinder is an ellipse*.

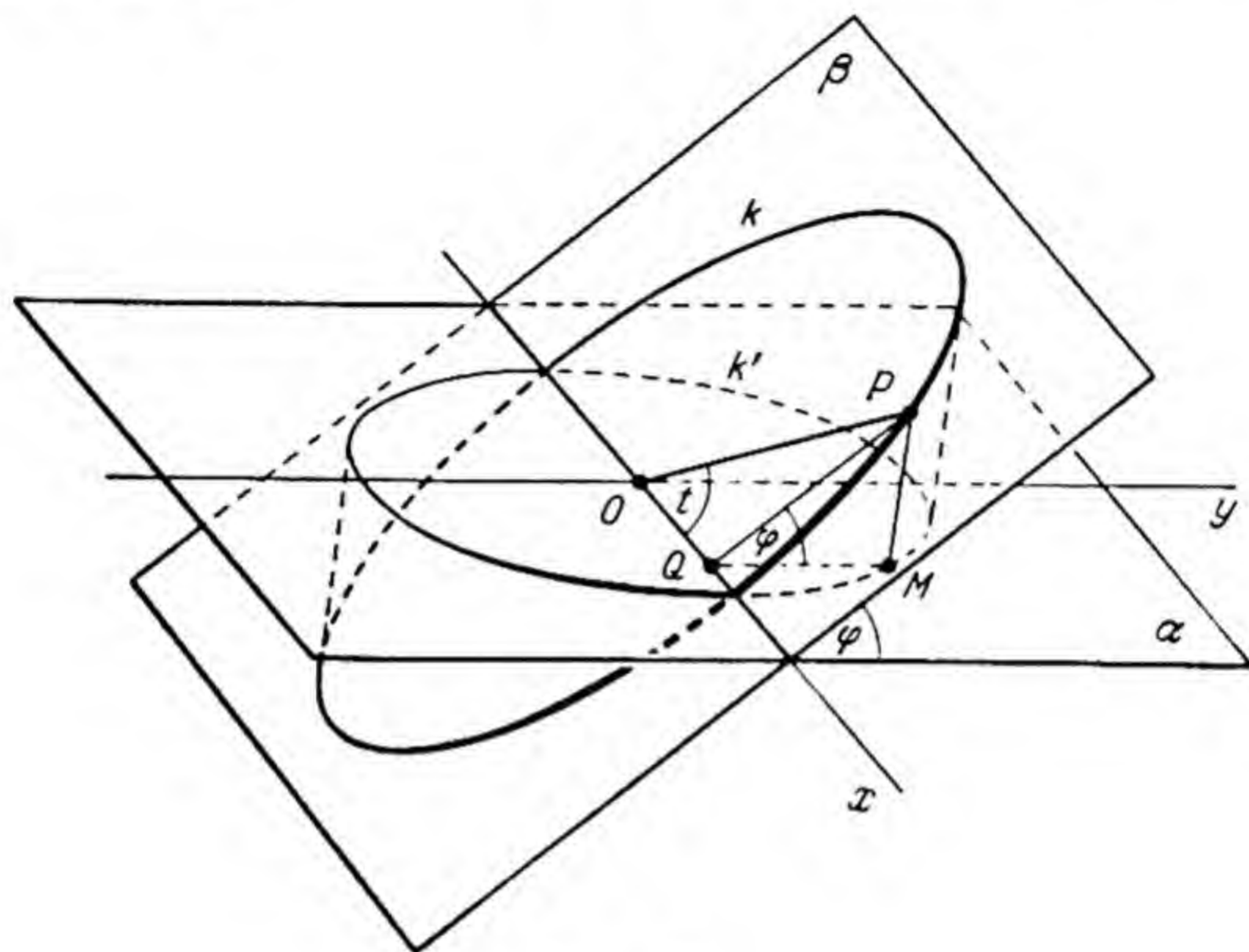


Fig. 50.

To prove this, consider a circular cylinder cut by a plane α (Fig. 51). Let k' be the curve formed by the intersection, and O the point where the plane α cuts the axis of the cylinder; let the plane β be passed through the point O perpendicular to the axis. The section of the cylinder made by this plane will be a circle k . Let the radius of this circle and the acute angle between the planes α and β be denoted by a and φ , respectively. Next, let us attach coordinate axes to the plane α , as shown in Fig. 51. Take an arbitrary point M on the curve k' ; let P be its projection on the plane β , Q its projection on the axis Ox , and t the angle which the segment OP makes with the axis Ox . Expressing the coordinates of the point M in terms of t , we obtain

$$x = OQ = OP \cdot \cos t = a \cos t,$$

$$y = QM = \frac{QP}{\cos \varphi} = \frac{OP \cdot \sin t}{\cos \varphi} = \frac{a}{\cos \varphi} \sin t.$$

Letting $\frac{a}{\cos \varphi} = b$, we get

$$x = a \cos t,$$

$$y = b \sin t.$$

These equations are the parametric equations of an ellipse; the curve k' is, therefore, an ellipse, as was to be shown.

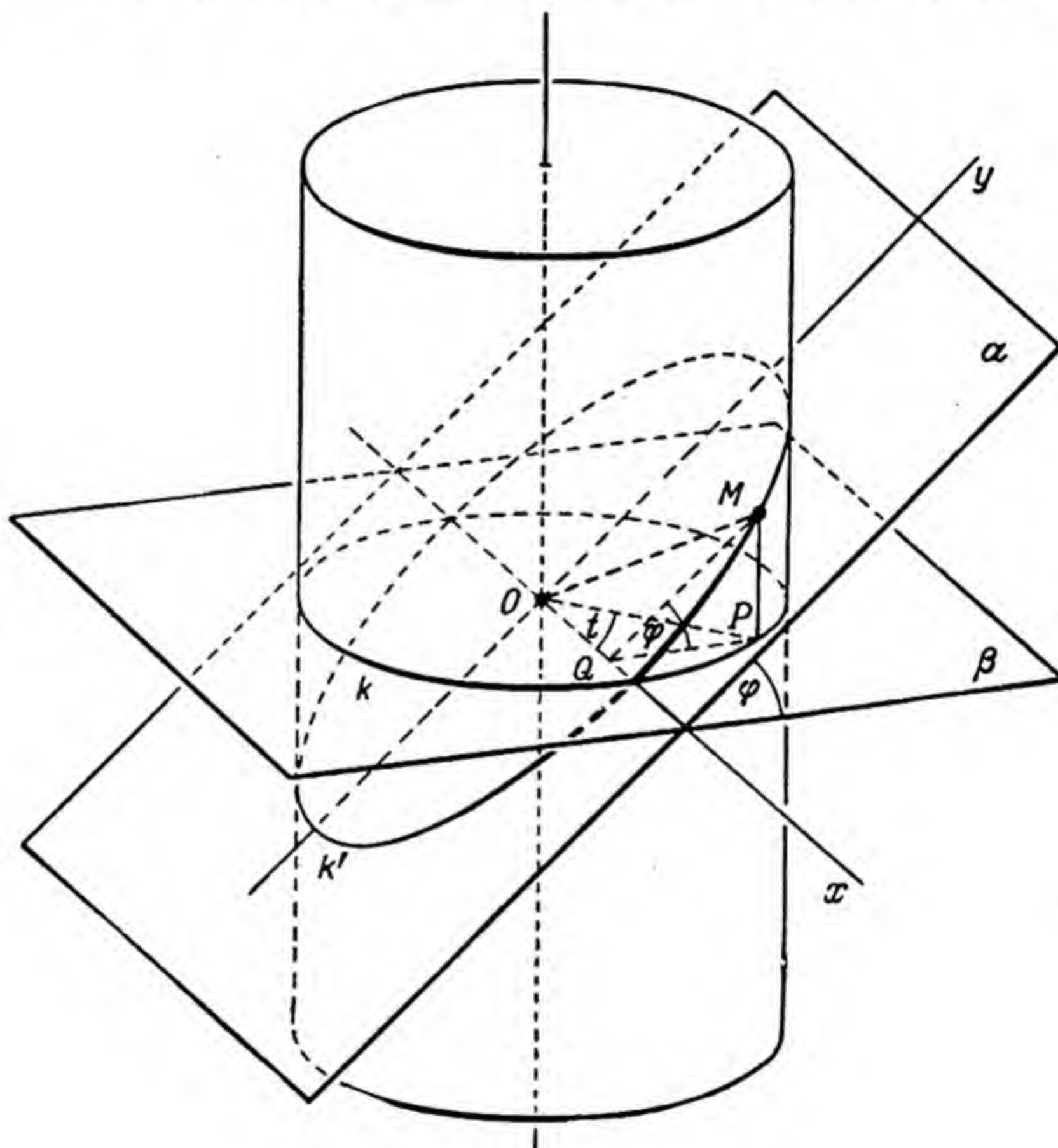


Fig. 51.

Note that $\frac{a}{\cos \varphi} > a$; hence a is the semi-minor axis of the ellipse k' , and $b = \frac{a}{\cos \varphi}$ is its semi-major axis, which means that the ellipse k' is elongated in the direction of the axis Oy .

The fact that the ellipse is the plane section of a circular cylinder, and also the projection of a circle on a plane, is very helpful in visualising the curve.

§ 30. The Hyperbola. Definition of the Hyperbola and Derivation of Its Canonical Equation

82. A hyperbola is the locus of points the difference of whose distances from two fixed points (called the foci) in the plane is numerically a constant; this constant is required to be less than the distance between the foci and different from zero. It is customary to denote the foci of a hyperbola by F_1 and F_2 , and the distance between them by $2c$.

Note. It is obvious that the difference of the distances of an arbitrary point M from two fixed points F_1 and F_2 cannot be greater than the distance between the points F_1 and F_2 . This difference is equal to the distance between F_1 and F_2 if, and only if, the point M lies on one of the extensions of the segment F_1F_2 . Consequently, the locus of points, the difference of whose distances from the two fixed points F_1, F_2 is a constant equal to the distance between F_1 and F_2 , consists of the two extensions of the segment F_1F_2 (Fig. 52).



Fig. 52.

If the difference of the distances of a point M from points F_1 and F_2 is equal to zero, then the point M is equidistant from

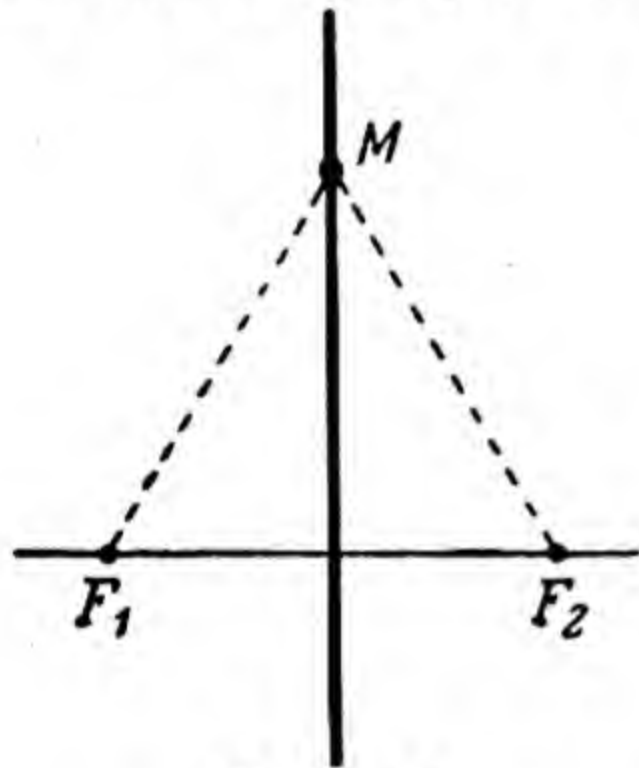


Fig. 53.

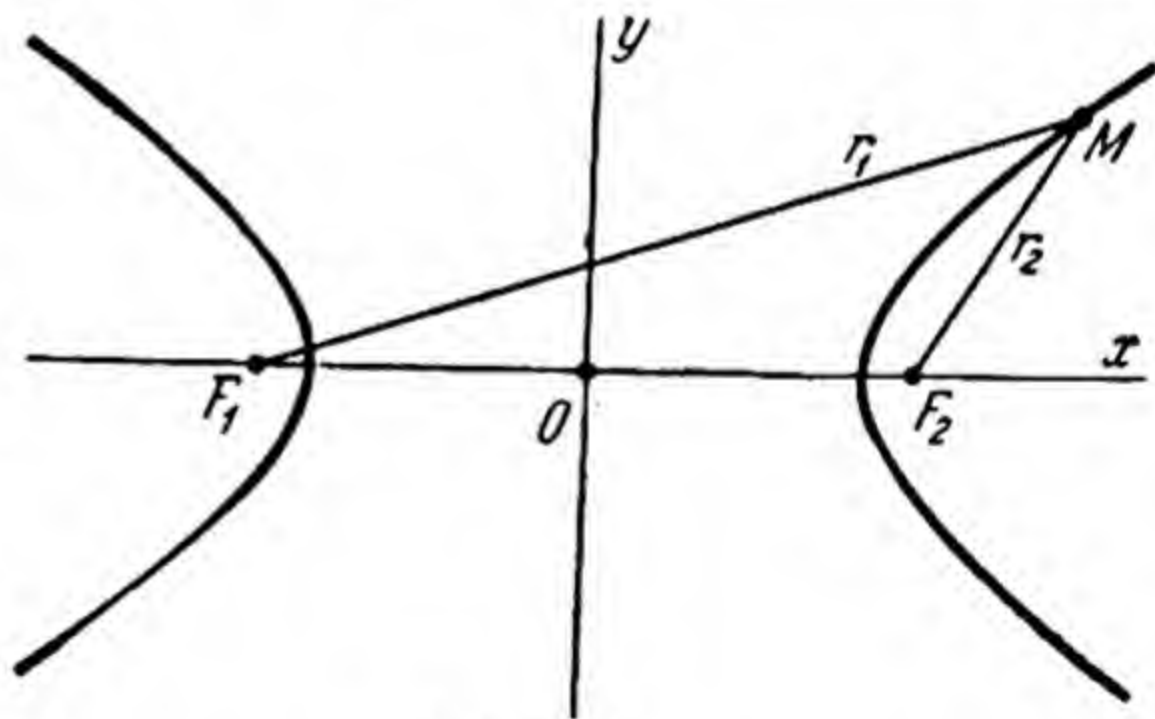


Fig. 54.

F_1 and F_2 . Hence, the locus of points, the difference of whose distances from the two fixed points F_1, F_2 is a constant equal to zero, is the perpendicular bisector of the segment F_1F_2 (Fig. 53).

These cases have been excluded by the restriction made at the end of the above definition.

83. Let M be an arbitrary point of a hyperbola with foci F_1 and F_2 (Fig. 54). The segments F_1M and F_2M (as well as the

lengths of these segments) are called the focal radii of the point M and are designated as r_1 and r_2 ($F_1M = r_1$, $F_2M = r_2$). By the definition of the hyperbola, the difference of the focal radii of its point M is a constant (that is, the difference of the focal radii of a point M remains the same for different positions of M on the hyperbola); this constant is generally denoted by $2a$. Thus, for any point M of the hyperbola, we have either

$$F_1M - F_2M = 2a, \quad (1)$$

when the point M is nearer to the focus F_2 , or

$$F_2M - F_1M = 2a, \quad (2)$$

when the point M is nearer to the focus F_1 .

By the definition of the hyperbola, $F_1M - F_2M < F_1F_2$ and $F_2M - F_1M < F_1F_2$; hence $2a < 2c$, that is,

$$a < c. \quad (3)$$

In the next article, we shall derive the equation of the hyperbola and then, by analysing this equation, establish the shape of the curve. We shall see that the hyperbola consists of two separate parts called the branches of the hyperbola, each branch extending indefinitely in two directions; the entire hyperbola is symmetric with respect to the line F_1F_2 , and also with respect to the perpendicular bisector of the segment F_1F_2 (see Fig. 54).

84. Let there be given a hyperbola with foci F_1, F_2 (we assume that a and b are also given). Let us attach to the plane a rectangular cartesian coordinate system, whose axes are specially chosen with respect to the hyperbola; namely, let the line F_1F_2 be taken as the x -axis, the direction from F_1 to F_2 adopted as positive, and the origin placed at the midpoint of the segment F_1F_2 (Fig. 54).

We proceed to derive the equation of the hyperbola referred to the chosen coordinate system. Take an arbitrary point M in the plane; designate its coordinates as x and y , and its focal radii F_1M and F_2M as r_1 and r_2 , respectively. The point M will lie on the (given) hyperbola if, and only if, $r_1 - r_2 = 2a$ or $r_2 - r_1 = 2a$. These two relations may be combined into

$$r_1 - r_2 = \pm 2a. \quad (4)$$

To obtain the desired equation of the hyperbola, it is necessary to express the variables r_1 and r_2 in terms of the current coordinates x, y and to substitute these expressions in (4). Since $F_1F_2 = 2c$, and since the foci F_1, F_2 are symmetrically situated on the axis Ox with respect to the origin, it follows that the coordi-

nates of the foci are $(-c, 0)$ and $(+c, 0)$, respectively; bearing this in mind and using formula (2) or Art. 18, we find

$$r_1 = \sqrt{(x+c)^2 + y^2}, \quad r_2 = \sqrt{(x-c)^2 + y^2}. \quad (5)$$

Substituting these expressions in (4), we obtain

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = \pm 2a. \quad (6)$$

The coordinates of a point $M(x, y)$ satisfy equations (6) if, and only if, the point M lies on the given hyperbola; consequently, (6) is the equation of this hyperbola referred to the chosen coordinate system (in fact, we have here two equations: one for the right-hand branch, and the other for the left-hand branch of the hyperbola).

The purpose of the remaining operations is to arrive at a simpler form of the equation of the hyperbola. Transposing the second radical to the right side of equation (6) and squaring both members, we obtain

$$(x+c)^2 + y^2 = 4a^2 \pm 4a\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2, \quad (7)$$

or

$$cx - a^2 = \pm a\sqrt{(x-c)^2 + y^2}. \quad (8)$$

Squaring both members of (7) yields

$$c^2x^2 - 2a^2cx + a^4 = a^2x^2 - 2a^2cx + a^2c^2 + a^2y^2, \quad (9)$$

whence

$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2). \quad (10)$$

We shall introduce here a new quantity

$$b = \sqrt{c^2 - a^2}; \quad (11)$$

the geometric meaning of the quantity b will be made clear a little later; we shall only note now that b is a real quantity (since, by Art. 83, $c > a$ and, consequently, $c^2 - a^2 > 0$). From (11), we have

$$b^2 = c^2 - a^2;$$

hence equation (10) may be written as

$$b^2x^2 - a^2y^2 = a^2b^2,$$

or

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (12)$$

Let us prove that *equation (12) is the equation of the given hyperbola*. This is not a self-evident fact since we have twice cleared radicals while reducing (6) to (12); it is obvious only that (12) is a consequence of (6).

We must prove that (6) is, in its turn, a consequence of (12), i. e., that these equations are equivalent.

Let x and y be any two numbers for which equation (12) is valid. Reversing the steps in the above derivation, we return from (12) to (10), and then to (9), which will now be written in the form

$$(cx - a^2)^2 = a^2 [(x - c)^2 + y^2].$$

Extracting the square root of both sides of this relation, we get

$$cx - a^2 = \pm a \sqrt{(x - c)^2 + y^2}. \quad (13)$$

If the point (x, y) is situated in the left half-plane, then $x < 0$ and the left member of (13) is negative. Consequently, in this case the right member of (13) must be taken with the minus sign. If, on the other hand, the point (x, y) lies in the right half-plane, then $x > 0$; from (12), we have $x \geq a$. Since $c > a$, it follows that $cx > a^2$, so that the left member of (13) is positive; in this case, the right member of (13) must therefore be taken with the plus sign. Thus, equation (13) has the same significance as equation (8). Next, on carrying out the necessary operations, we come from (8) back to (7), which will now be written as

$$(x + c)^2 + y^2 = [\sqrt{(x - c)^2 + y^2} \pm 2a]^2.$$

Hence

$$\sqrt{(x + c)^2 + y^2} = \pm [\sqrt{(x - c)^2 + y^2} \pm 2a]. \quad (14)$$

Let us determine which sign to choose *before the brackets* enclosing the right member of (14). We have to consider the following two cases:

(1) The point (x, y) lies in the *right* half-plane; then, according to the foregoing, the plus sign must be chosen within the brackets, the quantity enclosed in the brackets is positive, and the brackets must therefore be preceded by a plus sign.

(2) The point (x, y) is situated in the *left* half-plane. In this case, x is a negative number, so that the difference $x - c$ is numerically equal to the sum $|x| + c$. By (12), $|x| \geq a$; also, $c > a$. Therefore, $(x - c)^2 > 4a^2$, so that, of course, the sum $(x - c)^2 + y^2$ exceeds $4a^2$; hence, the square root of that sum is greater than $2a$, and the quantity enclosed in the brackets on the right side of (14) is, again, positive. Thus, in this case also, the

right member of (14) must be taken with the plus sign. We see that, for any position of the point (x, y) , equation (14) reduces to the form

$$\sqrt{(x+c)^2 + y^2} = \sqrt{(x-c)^2 + y^2} \pm 2a,$$

whence (6) can at once be obtained.

Thus, not only equation (12) is derivable from equation (6), but, conversely, (6) is derivable from (12). These equations are therefore equivalent, which proves that (12) is the equation of the given hyperbola.

Equation (12) is called *the canonical equation of the hyperbola*.

85. The equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

which represents the hyperbola in a certain system of rectangular cartesian coordinates is an equation of the second degree; accordingly, *the hyperbola is a curve of the second order*.

§ 31. Discussion of the Shape of the Hyperbola

86. We shall now investigate a hyperbola represented by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1. \quad (1)$$

Let us express the quantity y as a function of x by solving (1) for y :

$$y = \pm \sqrt{b^2 \left(\frac{x^2}{a^2} - 1 \right)},$$

or

$$y = \pm \frac{b}{a} \sqrt{x^2 - a^2}. \quad (2)$$

Since equation (1) contains only terms with even powers of the current coordinates x, y , it follows that the hyperbola represented by (1) is symmetric with respect to both coordinate axes (the proof of this is similar to that of the analogous assertion for the ellipse; see Art. 74); hence, it will clearly be sufficient to consider only the portion of the hyperbola in the first quadrant.

This portion of the hyperbola lies in the upper half-plane and is therefore associated with the plus sign before the right member of (2); at the same time, since the first-quadrant portion also lies

in the right half-plane, $x \geq 0$ for all its points. Thus, our task is to investigate the function

$$y = +\frac{b}{a}\sqrt{x^2 - a^2}, \quad (3)$$

where $x \geq 0$, and to draw the graph of this function.

To begin with, let $x = 0$. Substitution of $x = 0$ in the right member of (3) gives $y = \frac{b}{a}\sqrt{-a^2}$, that is, an imaginary number.

As x increases, y remains imaginary until x becomes equal to a . Setting $x = a$ in (3), we find $y = 0$. Consequently, the point $A(a, 0)$ is the extreme left point of the graph. As x increases further, y is continually real and positive in value; this is immediately evident from formula (3) since, for $x > a$, we have

$x^2 - a^2 > 0$. From formula (3), it is also evident that y is an increasing function of x (when $x \geq a$), which means that, as x increases, y does likewise all the time. Finally, we see from formula (3) that, as x increases indefinitely, y also increases indefinitely (as $x \rightarrow +\infty$, $y \rightarrow +\infty$). Bringing all these results together, we come to the following conclusion: As x increases from a , the variable point $M(x, y)$,

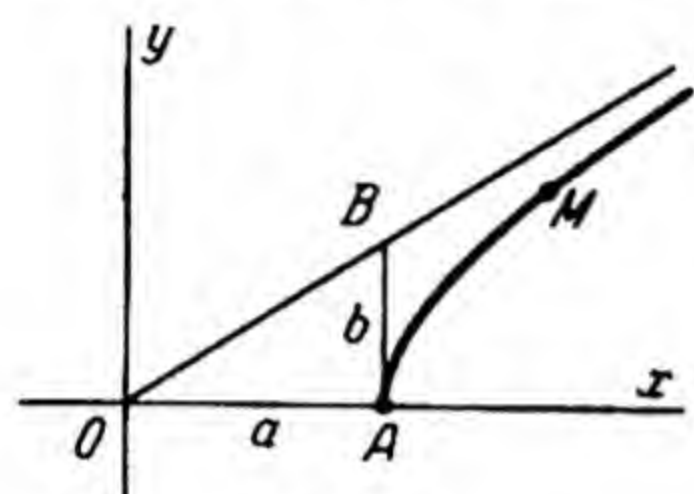


Fig. 55.

which describes the graph, moves continually to the "right" and "upwards", starting from the point $A(a, 0)$ and receding indefinitely from both the axis Oy (to the "right") and the axis Ox ("upwards"; see Fig. 55).

87. Let us examine more closely the manner in which the point M "goes to infinity". Besides the equation

$$y = +\frac{b}{a}\sqrt{x^2 - a^2}, \quad (4)$$

representing (for $x \geq a$) the portion of the hyperbola under investigation, let us consider for this purpose the equation

$$y = +\frac{b}{a}x, \quad (5)$$

which represents the straight line with slope $k = \frac{b}{a}$ and passing through the origin. Figure 55 shows that part of the line which lies in the first quadrant (to construct it, we have used the right triangle OAB with sides $OA = a$ and $AB = b$; obviously, the slope of the line OB is precisely $k = \frac{b}{a}$).

We shall now prove that, as the point M recedes to infinity, it approaches indefinitely close to the straight line $y = \frac{b}{a}x$.

Take an arbitrary value of $x (x \geq a)$ and consider two points: $M(x, y)$ and $N(x, Y)$, where

$$y = +\frac{b}{a}\sqrt{x^2 - a^2},$$

$$Y = \frac{b}{a}x.$$

The point $M(x, y)$ lies on the hyperbola (4), and the point $N(x, Y)$ lies on the line (5); since both points have the same abscissa x , the line joining the points M and N will be perpendicular to the axis Ox (Fig. 56). Let us compute the length of the segment MN .

First of all, note that

$$Y = +\frac{b}{a}x = \frac{b}{a}\sqrt{x^2} > +$$

$$+\frac{b}{a}\sqrt{x^2 - a^2} = y. \quad (6)$$

Hence $Y > y$ and, consequently, $MN = Y - y$. Now,

$$Y - y = \frac{b}{a}(x - \sqrt{x^2 - a^2}) =$$

$$= \frac{b}{a} \frac{(x - \sqrt{x^2 - a^2})(x + \sqrt{x^2 - a^2})}{x + \sqrt{x^2 - a^2}},$$

that is,

$$Y - y = \frac{ab}{x + \sqrt{x^2 - a^2}}. \quad (7)$$

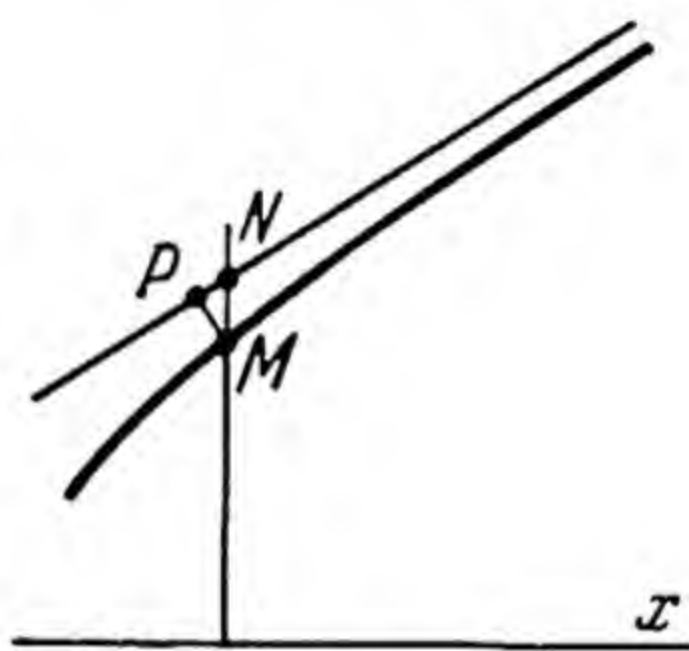


Fig. 56.

Let us analyse this last expression, assuming that $x \rightarrow +\infty$. Its denominator is the sum of two indefinitely increasing positive terms; as $x \rightarrow +\infty$, the denominator therefore tends to (positive) infinity. The numerator ab of our expression is a constant. Bringing these two facts together, we conclude that, as $x \rightarrow +\infty$, the right member of (7) tends to zero; hence $MN = Y - y$ also tends to zero.

Denote by P the foot of the perpendicular dropped from the point M to the line $y = \frac{b}{a}x$ (MP is thus the distance from M to that line). Obviously, $MP < MN$, and since $MN \rightarrow 0$, it follows that also $MP \rightarrow 0$, as was to be proved.

Thus, as the variable point M recedes to infinity along the first-quadrant portion of the hyperbola (1), the distance from the point M to the straight line $y = \frac{b}{a}x$ tends to zero.

88. Let Γ be an *arbitrary* curve, M a variable point on the curve, and let a be a straight line. If motion of the point M along the curve Γ is possible such that: (1) the point M recedes to infinity and, at the same time, (2) the distance from M to the line a tends to zero, then *the curve Γ is said to approach the line a asymptotically*. In this case, the line a is called *an asymptote* of the curve Γ .

Using the terms just introduced, we can formulate the result of the investigation carried out in Art. 87 as follows:

The graph of the function $y = \frac{b}{a} \sqrt{x^2 - a^2}$ (that is, the investigated portion of the hyperbola) approaches the line $y = \frac{b}{a} x$ asymptotically as $x \rightarrow +\infty$; or, the line $y = \frac{b}{a} x$ is an asymptote of the graph of the function $y = \frac{b}{a} \sqrt{x^2 - a^2}$ (and, at the same time, an asymptote of our hyperbola).

89. We shall now point out some further features of the position of the hyperbola relative to its asymptote (concerning ourselves, as before, only with the first-quadrant portion of the hyperbola).

Let us once more take the points $M(x, y)$ and $N(x, Y)$, considered in Art. 87, recalling that the point M lies on the hyperbola, and N on the asymptote. As has been established in Art. 87, $Y > y$. Hence, the point M is always "below" the point N . In other words, *the first-quadrant portion of the hyperbola (1) lies, throughout its extent, "below" its asymptote.*

Further, we have, according to formula (7),

$$Y - y = \frac{ab}{x + \sqrt{x^2 - a^2}}.$$

The denominator of the fraction is, for $x \geq a$, real and positive, and increases as x increases. Since the numerator is here a constant, it follows that the fraction continually decreases as x increases. Thus, it may be asserted that, as x *monotonically* tends to positive infinity (that is, continually increases), $MN = Y - y$ *monotonically* tends to zero (that is, continually decreases towards 0).

Let φ be the angle of inclination of the line $y = \frac{b}{a} x$, and let P be the foot of the perpendicular dropped from the point M to this line; then, obviously,

$$MP = MN \cdot \cos \varphi. \quad (8)$$

Since MN monotonically tends to zero, and since $\cos \varphi$ is a constant, it follows from (8) that MP also monotonically tends to zero.

In other words, whatever may be the position (in the first quadrant) of the point M on the hyperbola (4), as M moves along the hyperbola to the "right", its distance from the asymptote becomes continually smaller. We shall express this fact as follows: *The hyperbola approaches its asymptote monotonically.*

90. Let us summarise all that has been said in Arts. 86-89.

The first-quadrant portion of the hyperbola under investigation starts from the point $A(a, 0)$ and extends indefinitely out to the "right" and "upwards", asymptotically approaching the line $y = \frac{b}{a}x$ (from "below" and monotonically).

The graph in Fig. 55 has been drawn in accordance with the statement just formulated.

Note. The following two properties of our graph are also of importance: (1) its direction is perpendicular to Ox in the point

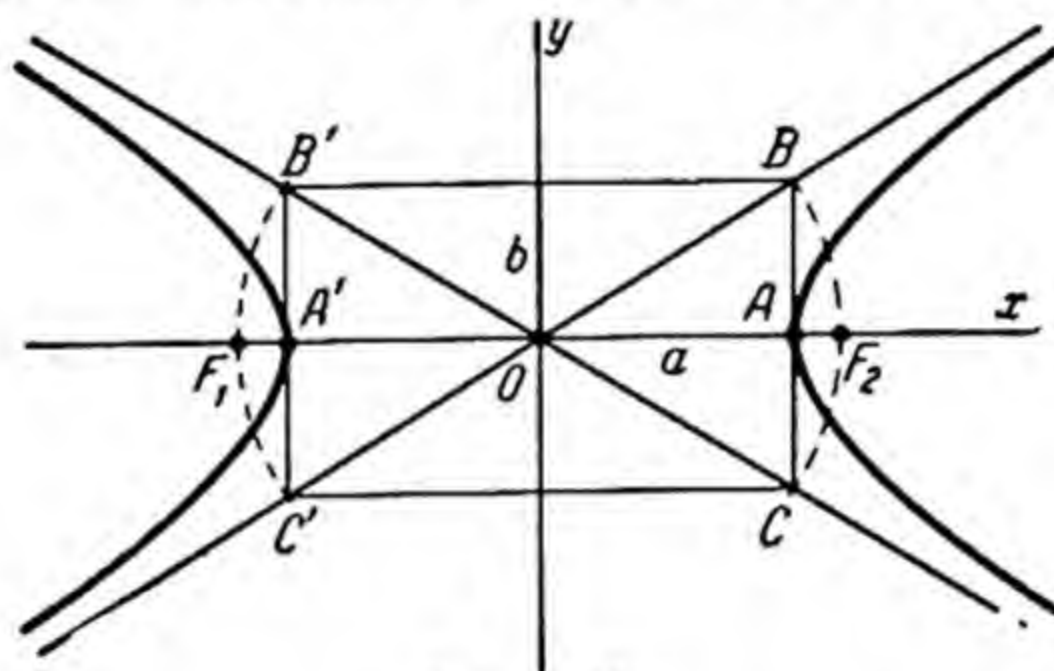


Fig. 57.

$A(a, 0)$; (2) it is convex "upwards" in all points. The proof of these properties will not, however, be given here, since the most natural methods for graph analysis of such kind are those furnished by the calculus.

91. Now that the portion of the hyperbola (4) in the first quadrant has been investigated, a general view of the entire hyperbola can be readily obtained by reflecting the graph in the coordinate axes.

The hyperbola represented by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

is shown in Fig. 57. It is easily seen that the entire hyperbola has two asymptotes,

$$y = \frac{b}{a}x$$

and

$$y = -\frac{b}{a}x;$$

we are already familiar with the first of these lines, and the second line is the reflection of the first in the axis Ox (or Oy).

Usually the axes of symmetry of a hyperbola are referred to simply as its *axes*, and the intersection of the axes is called *the centre* of the hyperbola. (In the case under consideration, the axes of the hyperbola are coincident with the coordinate axes.) One of the two axes (in our case, the one coincident with the axis Ox) intersects the hyperbola, whereas the other axis does not intersect it. The points of intersection of the hyperbola and the axis are called the *vertices*; a hyperbola has two vertices (marked by the letters A and A' in Fig. 57).

The rectangle with sides $2a$ and $2b$, which is symmetric with respect to the axes of a hyperbola and tangent to it at the vertices, is called *the fundamental rectangle* of the hyperbola (in Fig. 57, it is the rectangle $BB'C'C$). The diagonals of the fundamental rectangle of a hyperbola coincide with its asymptotes.

It should be noted that in mathematical books the term "axes of the hyperbola" is customarily applied also to the segments of lengths $2a$ and $2b$, joining the midpoints of the opposite sides of the fundamental rectangle. Accordingly, *the equation*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

is said to represent a hyperbola with semi-axes a and b .

Note. When making a sketch of a hyperbola with semi-axes a and b , one should begin by constructing the fundamental rectangle and then the asymptotes of the hyperbola. Following that, the hyperbola itself can be sketched in either "by eye", or after plotting a few of its points. In Fig. 57, it is shown (by dashed lines) how to locate the foci of a hyperbola by using its fundamental rectangle; this method is clearly based on the relation $c^2 = a^2 + b^2$, which follows from the formula (11) of Art. 84.

92. Consider now an equation of the form

$$-\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \quad (9)$$

By interchanging the letters x and y , a and b , this equation can be reduced to the equation studied in the preceding articles. Hence it is clear that equation (9) represents a hyperbola situated as shown in Fig. 58 (with its vertices B and B' lying on the axis

Oy). Equation (9) is also called the canonical equation of the hyperbola.

93. Two hyperbolas represented by the equations

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad -\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

in the same coordinate system and for the same values of a and b , are said to be *conjugate* to each other.

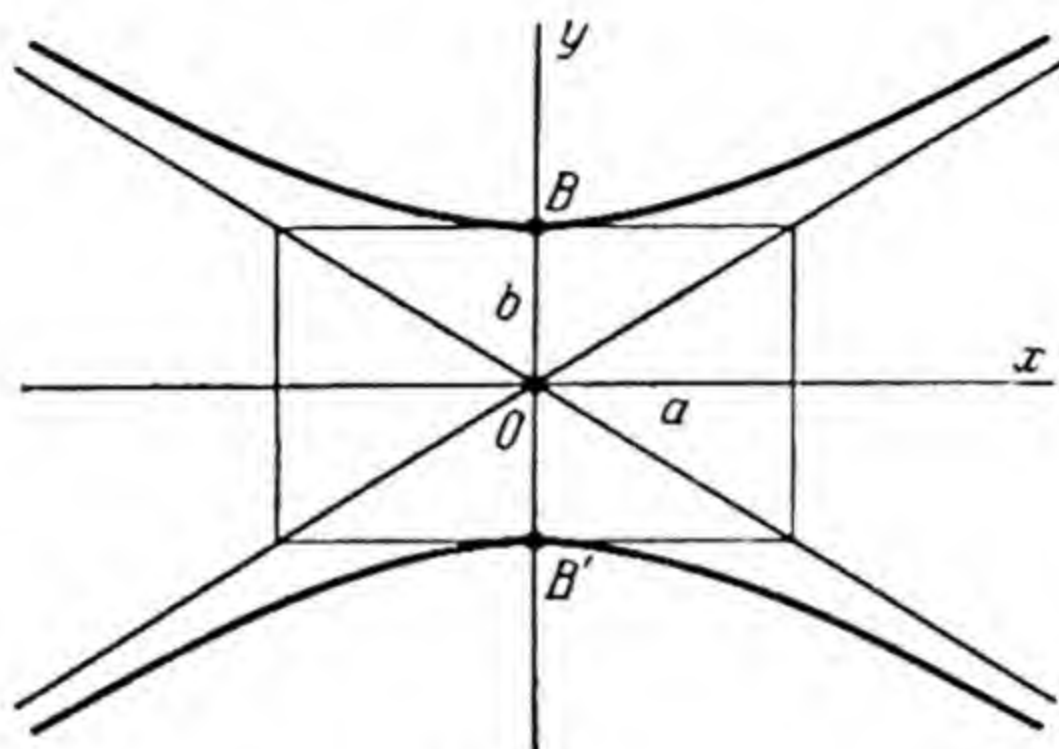


Fig. 58.

94. A hyperbola with equal semi-axes ($a = b$) is called an *equilateral hyperbola*. The canonical equation of an equilateral hyperbola may be written in the form

$$x^2 - y^2 = a^2.$$

The fundamental rectangle of an equilateral hyperbola is obviously a square; it is hence clear that *the asymptotes of an equilateral hyperbola are mutually perpendicular*.

§ 32. The Eccentricity of the Hyperbola

95. The *eccentricity* of a hyperbola is defined as the ratio of the distance between the foci of that hyperbola and the distance between its vertices; denoting the eccentricity by the letter ϵ , we obtain

$$\epsilon = \frac{c}{a}.$$

Since $c > a$ for a hyperbola, it follows that $\epsilon > 1$; that is, the eccentricity of every hyperbola is greater than unity.

Noting that $c^2 = a^2 + b^2$, we find

$$\epsilon^2 = \frac{c^2}{a^2} = \frac{a^2 + b^2}{a^2} = 1 + \left(\frac{b}{a}\right)^2,$$

whence

$$\epsilon = \sqrt{1 + \left(\frac{b}{a}\right)^2} \quad \text{and} \quad \frac{b}{a} = \sqrt{\epsilon^2 - 1}.$$

Accordingly, the eccentricity is determined by the ratio $\frac{b}{a}$ and, conversely, the ratio $\frac{b}{a}$ is determined by the eccentricity. Thus, *the eccentricity of a hyperbola characterises the shape of its fundamental rectangle and, hence, of the hyperbola itself.*

As the eccentricity decreases towards unity, $\epsilon^2 - 1$ decreases and, consequently, the ratio $\frac{b}{a}$ diminishes; this means that *the less the eccentricity of a hyperbola, the more elongated is its fundamental rectangle* (in the direction of the axis joining the vertices). In the case of an equilateral hyperbola, $a = b$ and $\epsilon = \sqrt{2}$.

§ 33. Rational Expressions for Focal Radii of the Hyperbola

96. Consider an arbitrary point $M(x, y)$ on a given hyperbola. If r_1 and r_2 are the focal radii of this point, then

$$r_1 = \sqrt{(x+c)^2 + y^2}, \quad r_2 = \sqrt{(x-c)^2 + y^2}. \quad (1)$$

However, focal radii may also be represented by formulas free of irrational terms. In fact, from equation (8) of Art. 84, we have

$$\sqrt{(x-c)^2 + y^2} = \pm \left(\frac{c}{a}x - a\right);$$

the plus sign refers here to the case when the point M is on the right-hand branch of the hyperbola. Letting $\frac{c}{a} = \epsilon$ and using the second of formulas (1) we obtain

$$r_2 = \pm (\epsilon x - a). \quad (2)$$

To express the first focal radius, we make use of the basic relation $r_1 - r_2 = \pm 2a$, where the plus sign likewise refers to points of the right-hand branch of the hyperbola. From this relation, we find $r_1 = r_2 \pm 2a = \pm (\epsilon x + a)$. Thus, for points of the right-hand branch of a hyperbola,

$$r_1 = \epsilon x + a, \quad r_2 = \epsilon x - a; \quad (3)$$

whereas for points of the left-hand branch,

$$r_1 = -(\varepsilon x + a), \quad r_2 = -(\varepsilon x - a). \quad (4)$$

These formulas will find an important application in the next section.

§ 34. The Directrices of the Ellipse and Hyperbola

97. Consider an ellipse in a rectangular cartesian coordinate system chosen so that the ellipse will be represented by the canonical equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

We assume that the ellipse is not a circle, i. e., that $a \neq b$ and, consequently, $e \neq 0$.

It is also assumed that the ellipse is elongated in the direction of the axis Ox ; i. e., that $a > b$.

The two straight lines perpendicular to the major axis of the ellipse and situated symmetrically with respect to the centre at a distance $\frac{a}{e}$ from it are called the directrices of the ellipse.

In the chosen coordinate system, the equations of the directrices are of the form

$$x = -\frac{a}{e} \quad \text{and} \quad x = +\frac{a}{e}.$$

We shall agree to refer to the first of the directrices as the left-hand directrix, and to the second as the right-hand directrix.

Since $e < 1$ for an ellipse, it follows that $\frac{a}{e} > a$. Hence, the right-hand directrix is situated to the right of the right-hand vertex of the ellipse; by analogy, the left-hand directrix lies to the left of the left-hand vertex. The ellipse together with the directrices is shown in Fig. 59.

98. Consider a hyperbola in a rectangular cartesian coordinate system chosen so that the hyperbola will be represented by the canonical equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

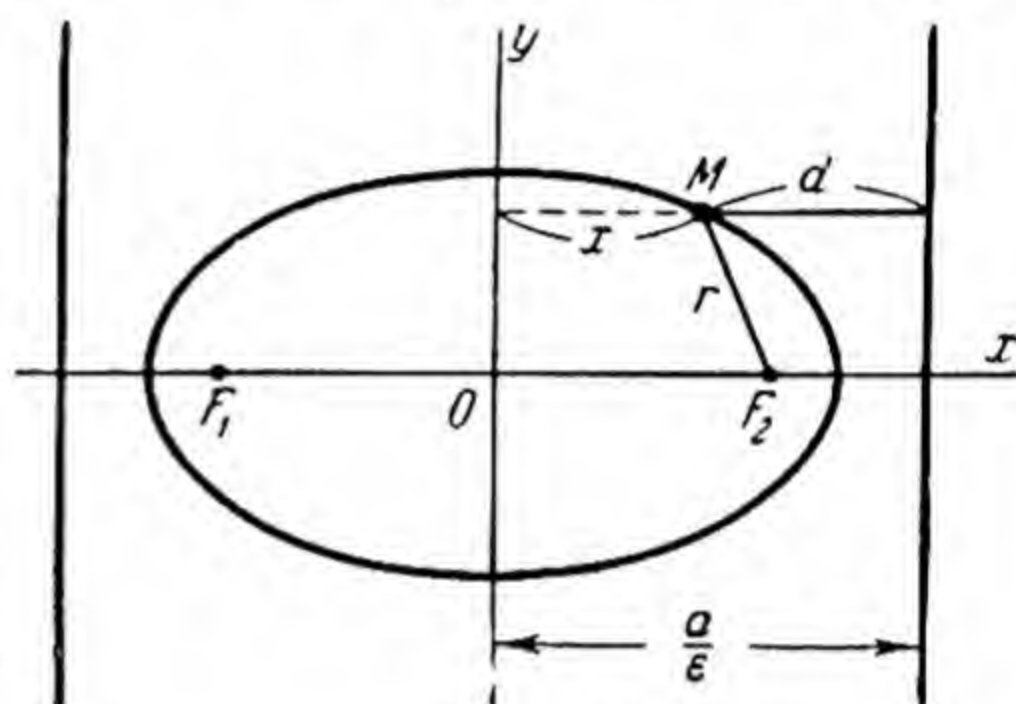


Fig. 59.

The two straight lines perpendicular to that axis of the hyperbola which intersects it, and situated symmetrically with respect to the centre at a distance $\frac{a}{\epsilon}$ from it, are called the *directrices* of the hyperbola.

In the chosen coordinate system, the equations of the directrices are of the form

$$x = -\frac{a}{\epsilon} \quad \text{and} \quad x = +\frac{a}{\epsilon}.$$

We shall agree to refer to the first of the directrices as the left-hand directrix, and to the second as the right-hand directrix.

Since $\epsilon > 1$ for a hyperbola, it follows that $\frac{a}{\epsilon} < a$. Hence the right-hand directrix is situated between the centre and the

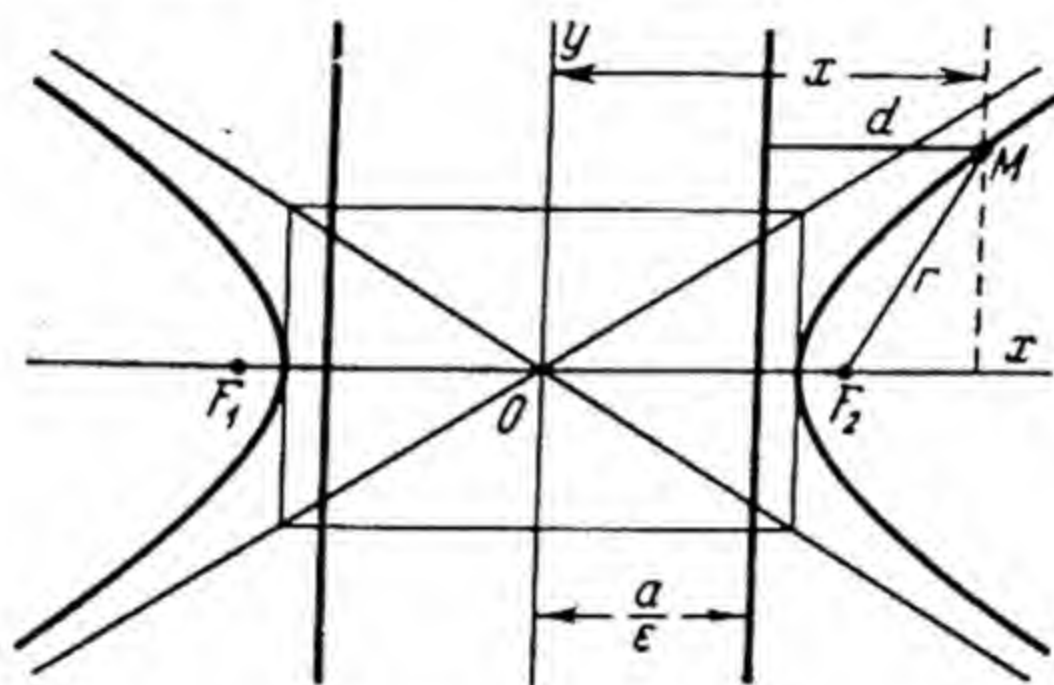


Fig. 60.

right-hand vertex of the hyperbola; by analogy, the left-hand directrix is situated between the centre and the left-hand vertex. Fig. 60 shows the hyperbola together with the directrices.

99. The meaning of the directrices of the ellipse and hyperbola is clarified by the following two theorems.

Theorem 11. *If r is the distance from an arbitrary point of an ellipse to one of its foci, and d the distance from the same point to the directrix corresponding to that focus, then the ratio $\frac{r}{d}$ is a constant equal to the eccentricity of the ellipse:*

$$\frac{r}{d} = \epsilon.$$

Proof. Let us take, for definiteness, the right-hand focus and the right-hand directrix. Let $M(x, y)$ be an arbitrary point of the

ellipse (see Fig. 59). The distance from M to the right-hand directrix is expressed by the relation

$$d = \frac{a}{\varepsilon} - x, \quad (1)$$

as is easily seen from the diagram; the distance from the point M to the right-hand focus is given by the second of formulas (2), § 27:

$$r = a - \varepsilon x. \quad (2)$$

From (1) and (2), we have

$$\frac{r}{d} = \frac{a - \varepsilon x}{\frac{a}{\varepsilon} - x} = \frac{(a - \varepsilon x) \varepsilon}{a - \varepsilon x} = \varepsilon.$$

The theorem is thus proved.

Theorem 12. *If r is the distance from an arbitrary point of a hyperbola to one of its foci, and d the distance from the same point to the directrix corresponding to that focus, then the ratio $\frac{r}{d}$ is a constant equal to the eccentricity of the hyperbola:*

$$\frac{r}{d} = \varepsilon.$$

Proof. Let us take, for definiteness, the right-hand focus and the right-hand directrix. Let $M(x, y)$ be an arbitrary point of the hyperbola (see Fig. 60). We have to consider the following two cases:

(1) The point M is on the right-hand branch of the hyperbola. Then the distance from M to the right-hand directrix is expressed by the relation

$$d = x - \frac{a}{\varepsilon}, \quad (3)$$

as is apparent from the diagram. The distance from the point M to the right-hand focus is given by the second of formulas (3), § 33:

$$r = \varepsilon x - a. \quad (4)$$

From (3) and (4), we have

$$\frac{r}{d} = \frac{\varepsilon x - a}{x - \frac{a}{\varepsilon}} = \frac{(\varepsilon x - a) \varepsilon}{\varepsilon x - a} = \varepsilon.$$

(2) The point M is on the left-hand branch of the hyperbola. Then the distance from M to the right-hand directrix is expressed by the relation

$$d = |x| + \frac{a}{\epsilon},$$

where $|x|$ is the distance from the point M to the axis Oy , $\frac{a}{\epsilon}$ the distance from the directrix to the axis Oy , and d the sum of these distances; but, since M lies on the left-hand branch of the hyperbola, x is a negative quantity, so that $|x| = -x$, and we obtain

$$d = -x + \frac{a}{\epsilon}. \quad (5)$$

The distance from M to the right-hand focus is given by the second of formulas (4), § 33:

$$r = -(\epsilon x - a). \quad (6)$$

From (5) and (6), we have

$$\frac{r}{d} = \frac{-(\epsilon x - a)}{-x + \frac{a}{\epsilon}} = \frac{(-\epsilon x + a)\epsilon}{-\epsilon x + a} = \epsilon.$$

The theorem is proved.

100. The property of the ellipse and hyperbola expressed by the above theorems serves as a basis for the following definition of these curves. *The locus of points whose distance r from a fixed point (a focus) is in a constant ratio*

$$\frac{r}{d} = \epsilon \quad (\epsilon = \text{const.})$$

to their distance d from a fixed straight line (the corresponding directrix), is an ellipse if $\epsilon < 1$, or a hyperbola if $\epsilon > 1$. (This statement can be verified by deriving the equation of this locus and ascertaining that the obtained equation is the equation of an ellipse or a hyperbola according as $\epsilon < 1$ or $\epsilon > 1$.)

The question naturally arises: What is the locus, defined in an analogous way for $\epsilon = 1$, that is, the locus of points, for each of which $r = d$? This locus turns out to be a new second-order curve, called the parabola.

§ 35. The Parabola. Derivation of the Canonical Equation of the Parabola

101. A parabola is the locus of points whose distance from a fixed point (called the focus) in the plane is equal to their distance from a fixed straight line (called the directrix and assumed not to pass through the focus).

It is customary to denote the focus of a parabola by the letter F , and the distance from the focus to the directrix by the letter p . The quantity p is called the *parameter* of a parabola. The curve is shown in Fig. 61 (the details of the drawing are fully explained in the next few articles).

Note. In accordance with Art. 100, a parabola is said to have eccentricity $\epsilon = 1$.

102. Let there be given a parabola (we assume that the parameter p is also given). Let us attach to the plane a rectangular cartesian coordinate system, whose axes are specially chosen with respect to the given parabola; namely, let the x -axis be drawn through the focus perpendicular to the directrix, the direction from the directrix to the focus adopted as positive on the x -axis, and the origin placed midway between the focus and the directrix (Fig. 61). We now proceed to derive the equation of the given parabola in this coordinate system.

Take an arbitrary point M in the plane and designate its coordinates as x and y . Let r denote the distance of the point M from the focus ($r = FM$), and d the distance of the point M from the directrix. The point M will lie on the given parabola if, and only if,

$$r = d. \quad (1)$$

In order to obtain the desired equation, it is necessary to express the variables r and d in terms of the current coordinates x , y and to substitute these expressions in (1). Note that the coordinates of the focus F are $(\frac{p}{2}, 0)$; bearing this in mind and using formula (2) of Art. 18, we find

$$r = \sqrt{\left(x - \frac{p}{2}\right)^2 + y^2}. \quad (2)$$

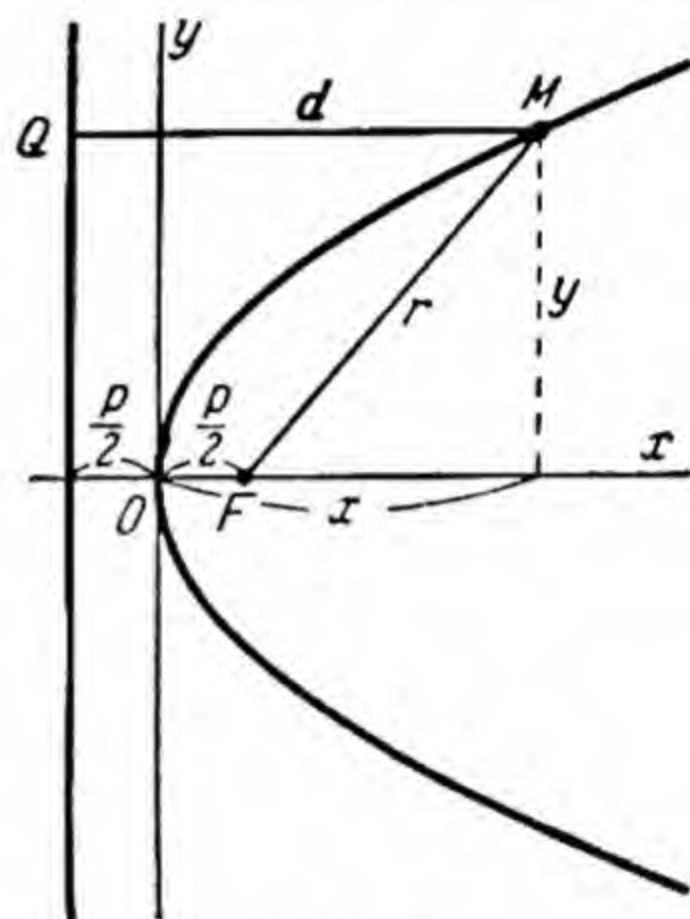


Fig. 61.

Denote by Q the foot of the perpendicular dropped from M upon the directrix. The coordinates of the point Q will clearly be $(-\frac{p}{2}, y)$; hence, by formula (2) of Art. 18, we obtain

$$d = MQ = \sqrt{\left(x + \frac{p}{2}\right)^2 + (y - y)^2} = x + \frac{p}{2} \quad (3)$$

(on extracting the root, we take $x + \frac{p}{2}$ with its original sign since $x + \frac{p}{2}$ is a positive number; this follows from the fact that the point $M(x, y)$ must lie on that side of the directrix where the focus is situated, that is, we must have $x > -\frac{p}{2}$, whence $x + \frac{p}{2} > 0$). Substituting expressions (2) and (3) for r and d in (1), we find

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = x + \frac{p}{2}. \quad (4)$$

The coordinates of a point $M(x, y)$ satisfy equation (4) if, and only if, the point M lies on the given parabola; accordingly, (4) is the equation of this parabola referred to the chosen coordinate system.

To reduce the equation of the parabola to a simpler form, we square both members of (4), which gives

$$x^2 - px + \frac{p^2}{4} + y^2 = x^2 + px + \frac{p^2}{4}, \quad (5)$$

or

$$y^2 = 2px. \quad (6)$$

We have derived equation (6) as a consequence of equation (4). It is easy to show that equation (4) may, in its turn, be derived as a consequence of (6). In fact, equation (5) is readily obtained from (6) by "retracing steps"; next, from (5) we get

$$\sqrt{\left(x - \frac{p}{2}\right)^2 + y^2} = \pm \left(x + \frac{p}{2}\right).$$

It remains to show that, if x, y satisfy equation (6), then the plus sign is here the only sign to choose. But this is clear since, from (6), $x = \frac{y^2}{2p}$ and, consequently, $x \geq 0$, so that $x + \frac{p}{2}$ is a positive number. Thus, we have come back to equation (4). Since each of equations (4) and (6) is a consequence of the other, they are equivalent. We hence conclude that *equation (6) is the equa-*

tion of the parabola. This equation is called the *canonical* equation of the parabola.

103. The equation $y^2 = 2px$, which represents the parabola in a certain system of rectangular cartesian coordinates is an equation of the second degree; accordingly, *the parabola is a curve of the second order*.

§ 36. Discussion of the Shape of the Parabola

104. Let us analyse the equation

$$y^2 = 2px \quad (1)$$

in order to form a clear idea of the shape of the parabola and thereby to show the correctness of its representation in Fig. 61.

Since equation (1) contains y only in an even power, the parabola represented by it is symmetrical with respect to the axis Ox . It will therefore be sufficient to investigate only the portion of the parabola which lies in the upper half-plane. This portion is represented by the equation

$$y = +\sqrt{2px}. \quad (2)$$

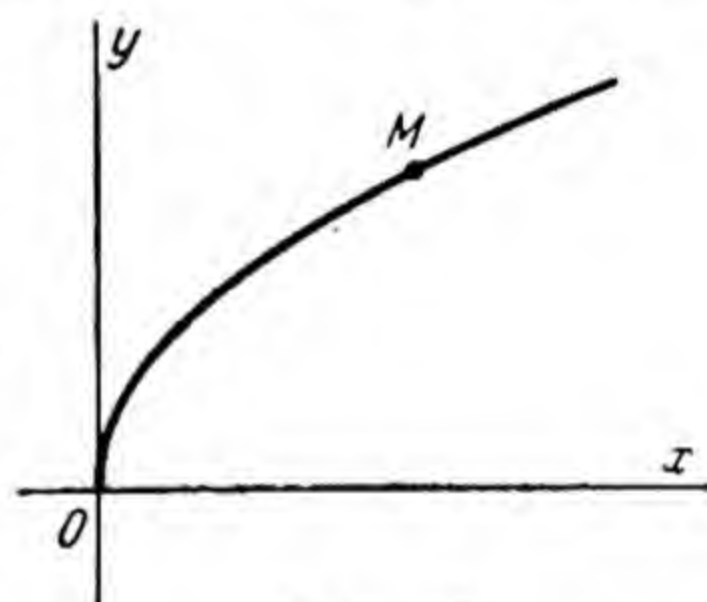


Fig. 62.

For negative values of x , equation (2) gives imaginary values of y . Consequently, no point of the parabola appears to the left of the axis Oy . For $x = 0$, we have $y = 0$. Hence the origin lies on the parabola and is its extreme "left" point. Equation (2) shows that, as x increases from zero, y continually increases. The equation also shows that, as $x \rightarrow +\infty$, $y \rightarrow +\infty$.

Thus, the variable point $M(x, y)$, which traces the portion of the parabola under consideration, moves to the "right" and "upwards", starting from the origin and receding indefinitely from both the axis Oy (to the "right") and the axis Ox ("upwards"; see Fig. 62).

Note. The following two properties of the parabola are also of importance: (1) the direction of the parabola is perpendicular to the axis Ox in the point $O(0, 0)$; (2) the portion of the parabola in the upper half-plane is convex "upwards". The graph in Fig. 62 has been drawn in accordance with these properties. Their proof will not, however, be given here, since the most natural methods for curve analysis of such kind are those furnished by the calculus.

105. Now that we have established the shape of the portion of the parabola lying in the upper half-plane, the determination of the shape of the entire parabola will present no difficulties; we have merely to reflect this portion of the curve in the axis Ox . The above-discussed Fig. 61 gives a general idea of the entire parabola represented by the equation

$$y^2 = 2px.$$

Usually the axis of symmetry of a parabola is referred to simply as its *axis* (in the case under consideration, the axis of the parabola coincides with the axis Ox). The point where a parabola cuts its axis is called *the vertex of the parabola* (in our case, the vertex is coincident with the origin). The number p , that is, the parameter of a parabola, represents the distance between the focus and the directrix. The geometric meaning of the parameter p may also be described as follows. Take some definite value of the abscissa, say $x = 1$, and find from equation (1) the corresponding values of the ordinate: $y = \pm \sqrt{2p}$. We obtain two points of the parabola, $M_1(1, +\sqrt{2p})$ and $M_2(1, -\sqrt{2p})$, symmetric with respect to the axis; the distance between these points is equal to $2\sqrt{2p}$. Thus, $2\sqrt{2p}$ is the length of the chord perpendicular to the axis and one unit of length distant from the vertex. We see that the length ($= 2\sqrt{2p}$) of this chord of the parabola increases with p . Consequently, the parameter p characterises the "spread" of a parabola, provided that this "spread" is measured perpendicular to the axis at a definite distance from the vertex.

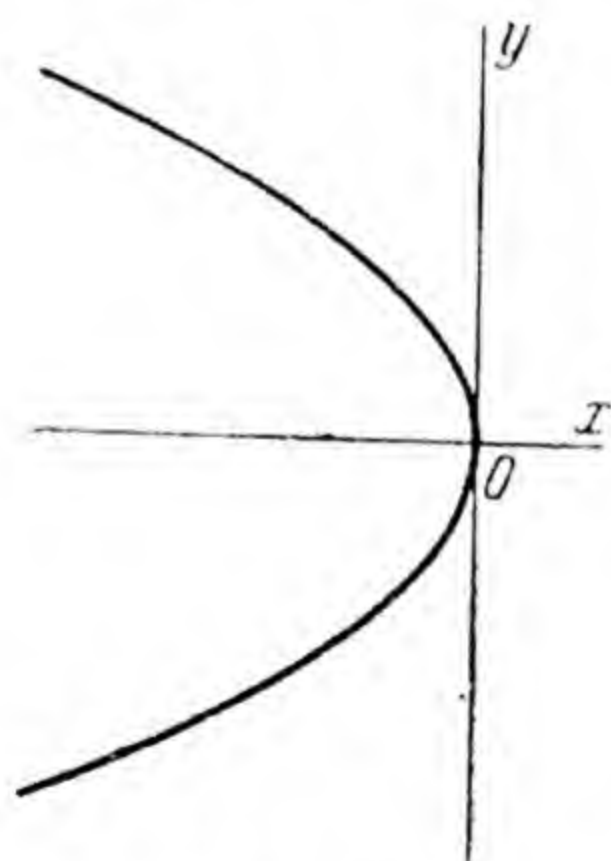


Fig. 63.

106. The equation

$$y^2 = -2px \quad (3)$$

(where p is positive) may be reduced to the equation $y^2 = 2px$ by substituting $-x$ for x , that is, by a transformation of coordinates corresponding to a reversal of the direction of the axis Ox . Hence, the equation $y^2 = -2px$ also represents a parabola whose axis is coincident with the axis Ox and whose vertex coincides with the origin; but this parabola is situated in the left half-plane, as shown in Fig. 63.

107. By analogy with the foregoing, we may assert that each of the equations

$$x^2 = 2py, \quad x^2 = -2py$$

(where $p > 0$) represents a parabola symmetric with respect to the axis Oy , with vertex at the origin (these equations, as well as equations (1) and (3), are referred to as the canonical equations of the parabola). A parabola represented by the equation $x^2 = 2py$ is said to open upwards; a parabola represented by

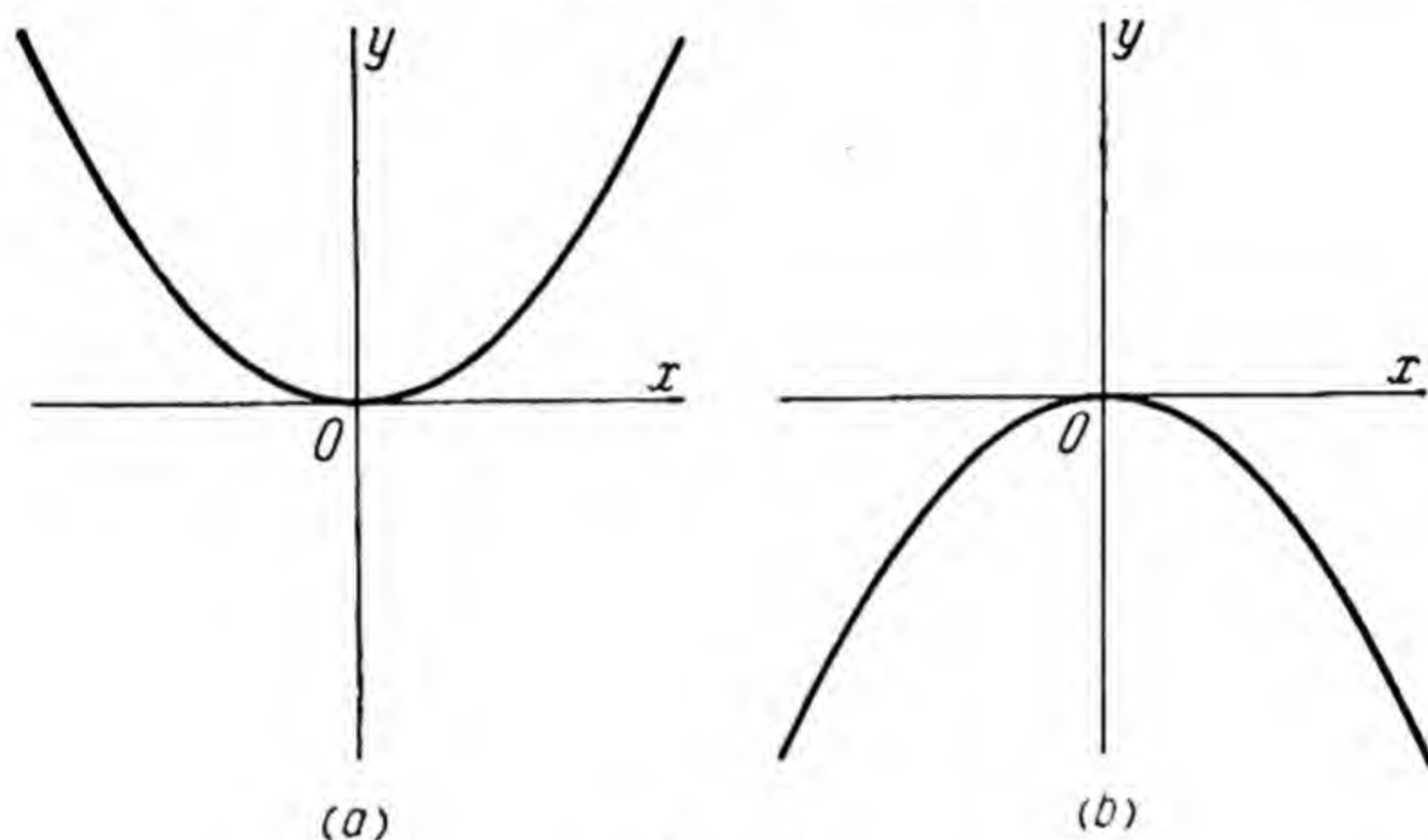


Fig. 64.

the equation $x^2 = -2py$ is said to open downwards (see Fig. 64 *a* and *b*, respectively); the use of these terms is natural and requires no further explanation.

§ 37. The Polar Equation of the Ellipse, Hyperbola and Parabola

108. Using the results of Arts 99-102, we shall now derive the polar equation of the ellipse, hyperbola and parabola (common in form to all the three curves) for a certain, specially chosen, position of the polar axis. It should be remarked, however, that in the case of the hyperbola this equation will represent only one of its branches, rather than the entire curve.

Let there be given any one of the above-mentioned curves: an ellipse, a hyperbola, or a parabola (if the given curve is a hyperbola, we shall consider a branch of it only). Denote the given curve by the letter L .

Let F be the focus of the curve, and g the directrix corresponding to that focus (in the case of a hyperbola, we shall denote by F and g the focus and directrix nearest to the branch under consideration).

Let us place the polar coordinate system so that the pole will coincide with the focus F , and the polar axis will be directed from the focus, along the axis of the curve L , away from the directrix g (Fig. 65). As usual, let ρ , θ denote the polar coordinates of a variable point M of the curve L . To derive the equation of the curve L , we shall use, as a basis, the relation

$$\frac{r}{d} = \varepsilon, \quad (1)$$

where ε is the eccentricity of the curve, and r and d have the same meaning as in Arts 99-102.

Since the pole is coincident with the focus F , it follows that

$$r = \rho. \quad (2)$$

Further,

$$d = QM = DN = DF + FN = DF + \rho \cos \theta. \quad (3)$$

Let P be the point whose position on the curve L is such that the line segment FP is perpendicular to the axis of L , and let p denote the length of the segment FP . In other words, p is equal to half the focal chord of L , perpendicular to the axis of the curve; the quantity p is called the focal parameter*) of the curve L .

From the basic relation (1), which refers to all points of the curve L , we have (for the point P , in particular)

$$\frac{FP}{SP} = \varepsilon,$$

whence $SP = \frac{FP}{\varepsilon} = \frac{p}{\varepsilon}$. Now, $SP = DF$, so that

$$DF = \frac{p}{\varepsilon}.$$

From this and from relation (3), we get

$$d = \frac{p}{\varepsilon} + \rho \cos \theta. \quad (4)$$

*) If the curve L is a parabola, $FP = PS$ (see Art. 101) and, consequently, $p = DF$, that is, p is equal to the distance between the focus and the directrix. In this case, therefore, the quantity p coincides with the already familiar parameter of the parabola, designated by the same letter.

Substituting expressions (2) and (4) for r and d in the left-hand member of (1), we find

$$\frac{\rho}{\frac{p}{\varepsilon} + \rho \cos \theta} = \varepsilon,$$

whence

$$\rho = \frac{p}{1 - \varepsilon \cos \theta}. \quad (5)$$

This is the polar equation of the ellipse, the hyperbola (or rather, one branch of the hyperbola) and the parabola. Here p is the local parameter, and ε the eccentricity of the curve. Equation (5) is used in mechanics.

§ 38. Diameters of Curves of the Second Order

109. An important and, at first glance, surprising property of second-order curves (ellipses, hyperbolas and parabolas) is expressed by the following

Theorem 13. *The midpoints of parallel chords of a second-order curve lie on a straight line.*

Proof. (1) Let the given curve be the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (1)$$

(Fig. 66). Denote by k the common slope of parallel chords; then the equation of each of them may be written as

$$y = kx + l, \quad (2)$$

where l has different values for different chords. Let us find the endpoints of a chord represented by equation (2) for some value of l . Solving (1) and (2) simultaneously and eliminating y from them, we obtain

$$\frac{x^2}{a^2} + \frac{(kx + l)^2}{b^2} = 1,$$

or

$$(b^2 + a^2k^2)x^2 + 2a^2klx + a^2(l^2 - b^2) = 0. \quad (3)$$

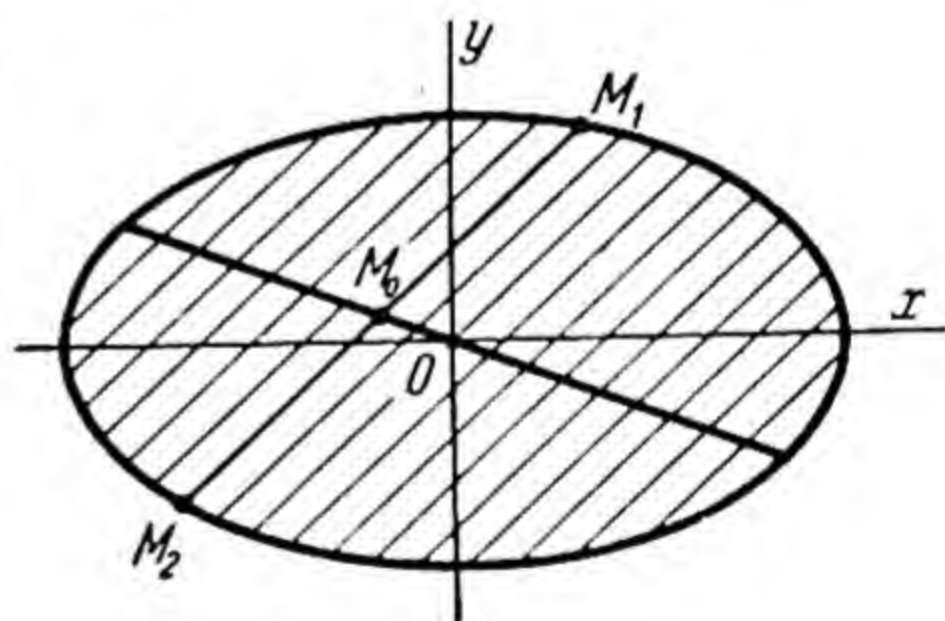


Fig. 66.

The roots x_1, x_2 of this quadratic are the abscissas of the endpoints M_1, M_2 of the chord. Let $M_0(x_0, y_0)$ be the midpoint of the chord; then

$$x_0 = \frac{x_1 + x_2}{2}.$$

Now, by the well-known theorem concerning the sum of the roots of a quadratic equation,

$$x_1 + x_2 = -\frac{2a^2kl}{b^2 + a^2k^2}.$$

Hence

$$x_0 = -\frac{a^2 kl}{b^2 + a^2 k^2}.$$

Having found x_0 , we find y_0 from (2):

$$y_0 = kx_0 + l = -\frac{a^2 k^2 l}{b^2 + a^2 k^2} + l = \frac{b^2 l}{b^2 + a^2 k^2}.$$

Thus,

$$x_0 = -\frac{a^2 kl}{b^2 + a^2 k^2}, \quad y_0 = \frac{b^2 l}{b^2 + a^2 k^2}. \quad (4)$$

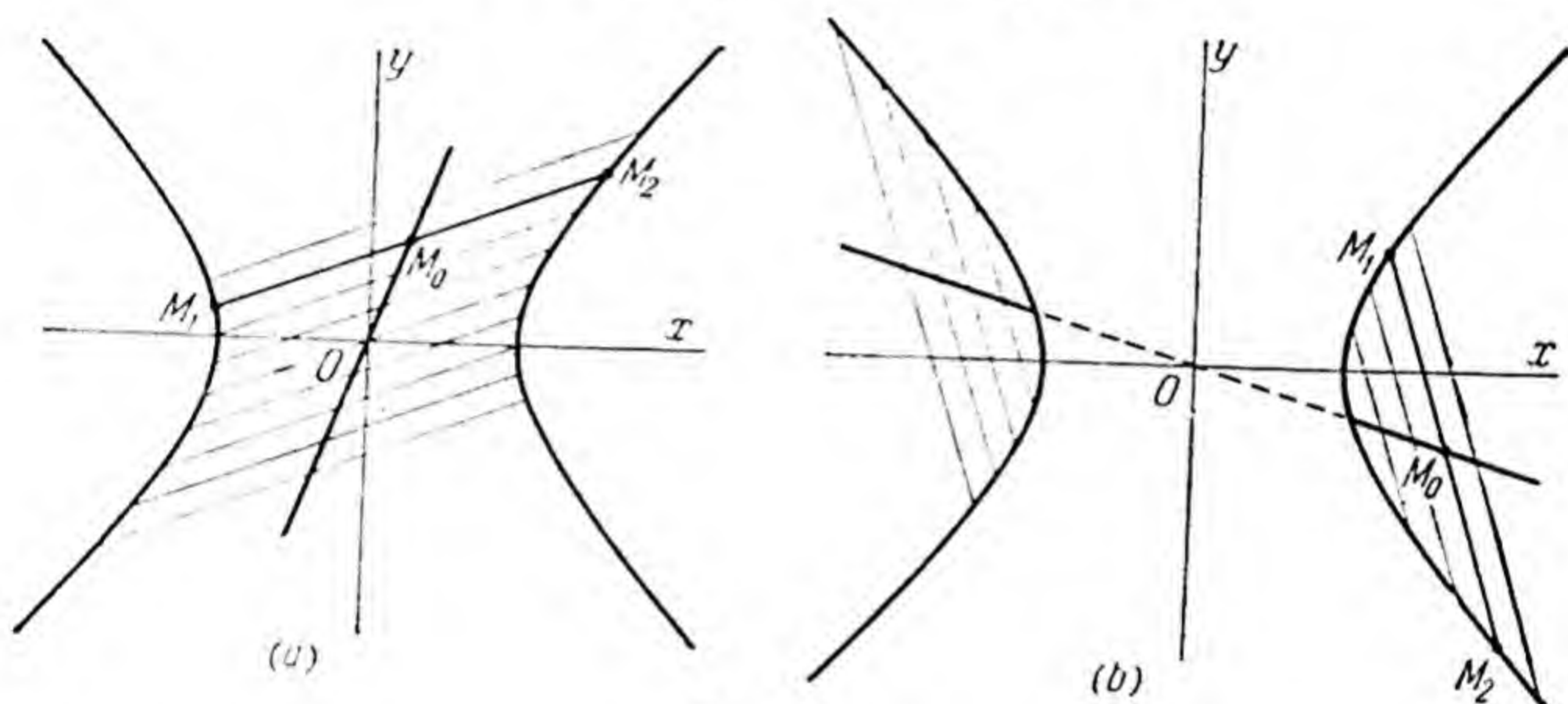


Fig. 67.

By varying here the value of l , we shall obtain the coordinates x_0, y_0 of the midpoints of different parallel chords of the ellipse; but, as is clear from relations (4), x_0 and y_0 will invariably be connected by the equation

$$\frac{y_0}{x_0} = -\frac{b^2}{a^2 k},$$

or $y_0 = k'x_0$, where

$$k' = -\frac{b^2}{a^2 k}. \quad (5)$$

Thus, the midpoints of all chords of slope k lie on the straight line

$$y = k'x. \quad (6)$$

(2) Let the given curve be the *hyperbola*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad (7)$$

(Fig. 67a and b). Denote by k the common slope of parallel chords; then the equation of each of them may be written as

$$y = kx + l. \quad (8)$$

Before proceeding further, note that no chords of a hyperbola can be parallel to its asymptotes (since every line parallel to an asymptote meets the hyperbola at one point only); therefore, $k \neq \frac{b}{a}$ and $k \neq -\frac{b}{a}$. Let us find the end-points of a chord represented by equation (8) for some value of l . Eliminating y between (7) and (8), we get

$$\frac{x^2}{a^2} - \frac{(kx + l)^2}{b^2} = 1,$$

or

$$(b^2 - a^2k^2)x^2 - 2a^2klx - a^2(l^2 + b^2) = 0. \quad (9)$$

Since $k \neq \pm \frac{b}{a}$, it follows that $b^2 - a^2k^2 \neq 0$. Consequently, (9) is a quadratic equation. The roots x_1, x_2 of this quadratic are the abscissas of the end-points M_1, M_2 of the chord. Let $M_0(x_0, y_0)$ be the midpoint of the chord; then

$$x_0 = \frac{x_1 + x_2}{2}.$$

Using the theorem on the sum of the roots of a quadratic equation, we find

$$x_1 + x_2 = \frac{2a^2kl}{b^2 - a^2k^2}.$$

Hence, $x_0 = \frac{a^2kl}{b^2 - a^2k^2}$. Now that x_0 is known, we find y_0 from (8):

$$y_0 = kx_0 + l = \frac{a^2k^2l}{b^2 - a^2k^2} + l = \frac{b^2l}{b^2 - a^2k^2}.$$

Thus,

$$x_0 = \frac{a^2kl}{b^2 - a^2k^2}, \quad y_0 = \frac{b^2l}{b^2 - a^2k^2}. \quad (10)$$

By varying here the value of l , we shall obtain the coordinates x_0, y_0 of the midpoints of different parallel chords of the hyperbola; but, as is clear from relations (10), x_0 and y_0 will invariably be connected by the equation

$$\frac{y_0}{x_0} = \frac{b^2}{a^2k},$$

or $y_0 = k'x_0$, where

$$k' = \frac{b^2}{a^2k}. \quad (11)$$

Thus, the midpoints of all chords of slope k lie on the straight line

$$y = k'x. \quad (12)$$

(3) Finally, let the given curve be the *parabola*

$$y^2 = 2px \quad (13)$$

(Fig. 68). Denote by k the common slope of parallel chords; then the equation of each of them may be written as

$$y = kx + l. \quad (14)$$

Before proceeding further, note that no chords of a parabola can be parallel to its axis (since every straight line parallel to the axis meets the parabola at one point only); therefore $k \neq 0$.

Let us find the endpoints of a chord represented by equation (14) for some value of l . Eliminating y between (13) and (14), we get

$$(kx + l)^2 - 2px = 0,$$

or

$$k^2x^2 + 2(kl - p)x + l^2 = 0. \quad (15)$$

Since $k \neq 0$, it follows that (15) is a quadratic equation. The roots x_1, x_2 of this equation are the abscissas of the endpoints M_1, M_2 of the chord. Let $M_0(x_0, y_0)$ be the midpoint of this chord; then we have

$$x_0 = \frac{x_1 + x_2}{2};$$

by the theorem on the sum of the roots of a quadratic,

$$x_1 + x_2 = -\frac{2(kl - p)}{k^2}.$$

Consequently, $x_0 = \frac{p - kl}{k^2}$. Now that x_0 is known, we find y_0 from (14):

$$y_0 = kx_0 + l = k \frac{p - kl}{k^2} + l = \frac{p}{k}.$$

Thus,

$$x_0 = \frac{p - kl}{k^2}, \quad y_0 = \frac{p}{k}. \quad (16)$$

By varying here the value of l , we shall get the coordinates of the midpoints of different parallel chords of the parabola; but, as is clear from relations (16), y_0 will invariably be equal to the number $\frac{p}{k}$. Thus, the midpoints of all chords of slope k lie on the straight line

$$y = \frac{p}{k}, \quad (17)$$

which is parallel to the x -axis and also to the axis of the parabola.

We could now regard the theorem as completely proved, were it not for a certain defect in our computation technique. Namely, we represented chords of a second-order curve by an equation in the slope-intercept form $y = kx + l$. Our computations, therefore, must become meaningless when the chords under consideration are parallel to the axis Oy (since straight lines parallel to the axis Oy have no slope). For such curves, however, the validity of the theorem follows at once from the symmetric properties of the ellipse, hyperbola and parabola. For, the ellipse, hyperbola and parabola represented by the canonical equations (1), (7) and (13) are symmetric with respect to the axis Ox . Consequently, when the chords of these curves are parallel to the axis Oy , their midpoints still lie on a straight line (in this case, on the axis Ox).

110. The straight line passing through the midpoints of parallel chords of a second-order curve is called a diameter of that curve.

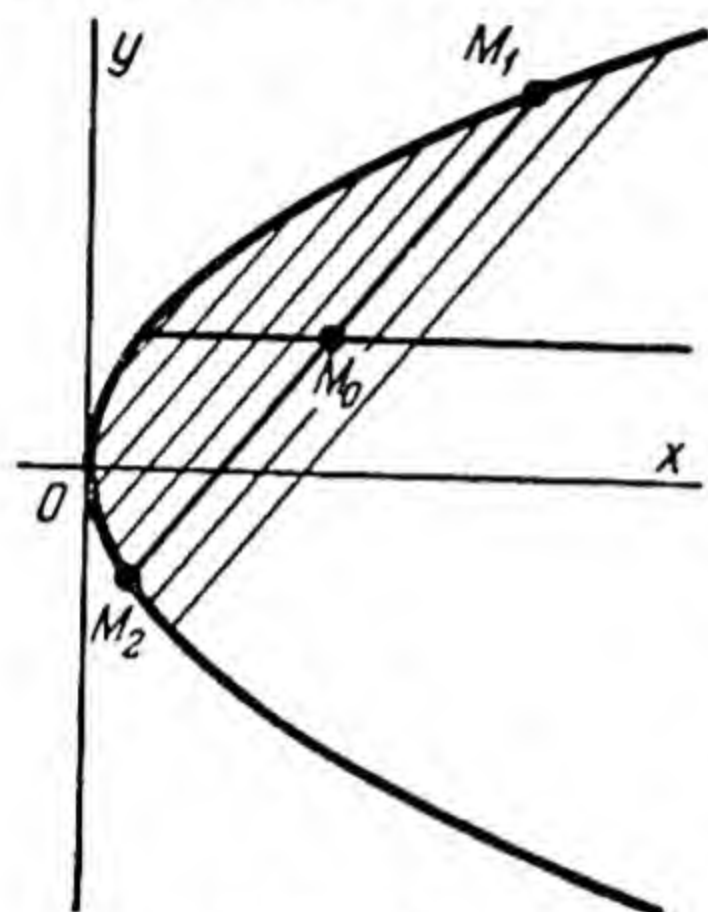


Fig. 68.

All diameters of an ellipse or a hyperbola pass through the centre of the curve; this is clear geometrically (since the centre is the midpoint of every chord passing through it), and is also immediately evident from equations (6) and (12) of Art. 109.

According to equation (17), all diameters of a parabola are parallel to its axis.

We shall now point out some properties of the diameters of the ellipse and hyperbola.

Consider the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Let k be the slope of a diameter of the ellipse. Draw chords parallel to this diameter; the locus of their midpoints is a second diameter, which is said to be *conjugate* to the first. The slope k' of the second diameter is determined from (5), which gives

$$kk' = -\frac{b^2}{a^2}. \quad (18)$$

Let us now find the diameter conjugate to the diameter of slope k' ; analogous to the above, the slope k'' of this new diameter will be determined by the relation

$$k'k'' = -\frac{b^2}{a^2}.$$

Hence, from (18), we obtain: $k'' = k$.

Thus, if one diameter of an ellipse is conjugate to another diameter, then this second diameter is conjugate to the first. Such diameters are therefore called *conjugate diameters*. Relation (18) is referred to as the condition that the diameters (of an ellipse) of slopes k and k' should be conjugate.

The reciprocity of conjugate diameters may also be expressed as follows: If one diameter of an ellipse bisects the chords parallel to another diameter, then this second diameter bisects the chords parallel to the first (Fig. 69; this diagram also illustrates an interesting consequence of the foregoing proposition, namely, that the tangent lines to an ellipse at the ends of its diameter are parallel to each other and to the conjugate diameter).

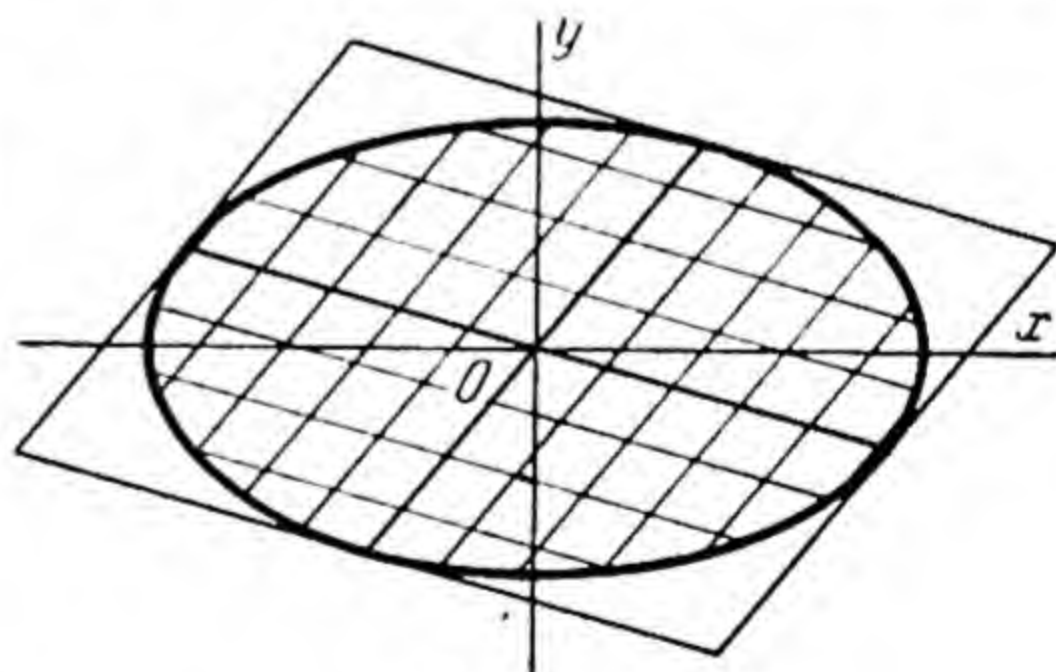


Fig. 69.

All that has just been said about the diameters of the ellipse, is directly applicable to the diameters of the hyperbola, except that the condition for conjugate diameters of the hyperbola is somewhat different from (18). In fact, for a hyperbola represented by the equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

the condition that its diameters of slopes k and k' should be conjugate is

$$kk' = -\frac{b^2}{a^2}, \quad (19)$$

as follows from relation (11).

Note. The axes of symmetry of an ellipse (and of a hyperbola) form a pair of conjugate diameters, since each axis bisects the chords parallel to the other. The axes of symmetry differ from all other pairs of conjugate diameters in being mutually perpendicular as well as conjugate.

§ 39. The Optical Properties of the Ellipse, Hyperbola and Parabola

111. Among the most remarkable properties of the ellipse, hyperbola and parabola are their so-called optical properties. Incidentally, these properties show that the term "foci" owes its origin to physics.

Let us, first of all, formulate these properties from a purely geometric viewpoint.

1. The line tangent to an ellipse at a point M makes equal angles with the focal radii F_1M , F_2M and passes externally to the angle F_1MF_2 (Fig. 70a).

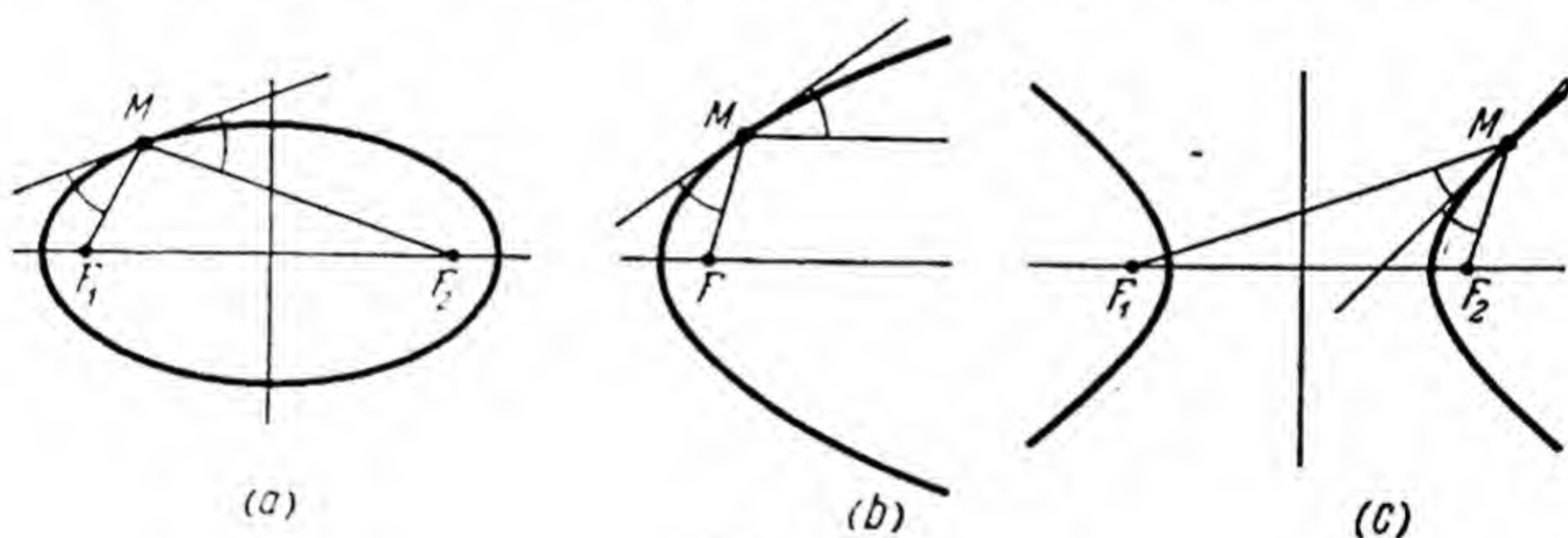


Fig. 70.

2. The line tangent to a parabola at a point M makes equal angles with the focal radius FM and the ray drawn from M parallel to the axis of the parabola, in the direction in which the parabola opens (Fig. 70b).

3. The line tangent to a hyperbola at a point M makes equal angles with the focal radii F_1M , F_2M and passes within the angle F_1MF_2 (Fig. 70c).

We shall omit the proof of these properties. It will be sufficient here to note that, in order to prove them analytically, one must be able to find the slope of a tangent line, given the equation of the curve and the point of tangency. The appropriate rules are given in textbooks on the calculus. To clarify the physical meaning of the above propositions, let us imagine that an ellipse, parabola or hyperbola is revolved about its axis (containing the foci), thereby generating a surface called an ellipsoid, paraboloid or hyperboloid, respectively. A physical surface of such shape, when silvered, will form an elliptic, parabolic or hyperbolic mirror, respectively. Recalling the optical laws of reflection, we conclude that:

1. If a source of light is placed at one of the foci of an elliptic mirror, its rays will, after reflection at the mirror, converge to the other focus,

2. If a source of light is placed at the focus of a parabolic mirror, its rays will, after reflection at the mirror, be parallel to the axis.

3. If a source of light is placed at one of the foci of a hyperbolic mirror, its rays will, after reflection at the mirror, appear to emanate from the other focus.

The above property of a parabolic mirror is utilised in searchlights.

§ 40. The Ellipse, Hyperbola and Parabola as Conic Sections

112. A new light is cast on the geometric nature of ellipses, hyperbolas and parabolas by the following

Theorem 14. *A section of any circular cone made by a plane (not passing through the vertex of the cone) is a curve no other than an ellipse, hyperbola, or parabola. If the cutting plane cuts only one nappe of the cone and the intersection is a closed curve, this curve is an ellipse; if the plane cuts only one nappe of the cone and the intersection is an open curve, this curve is a parabola; if the plane cuts both nappes of the cone, the section is a hyperbola (Fig. 71).*

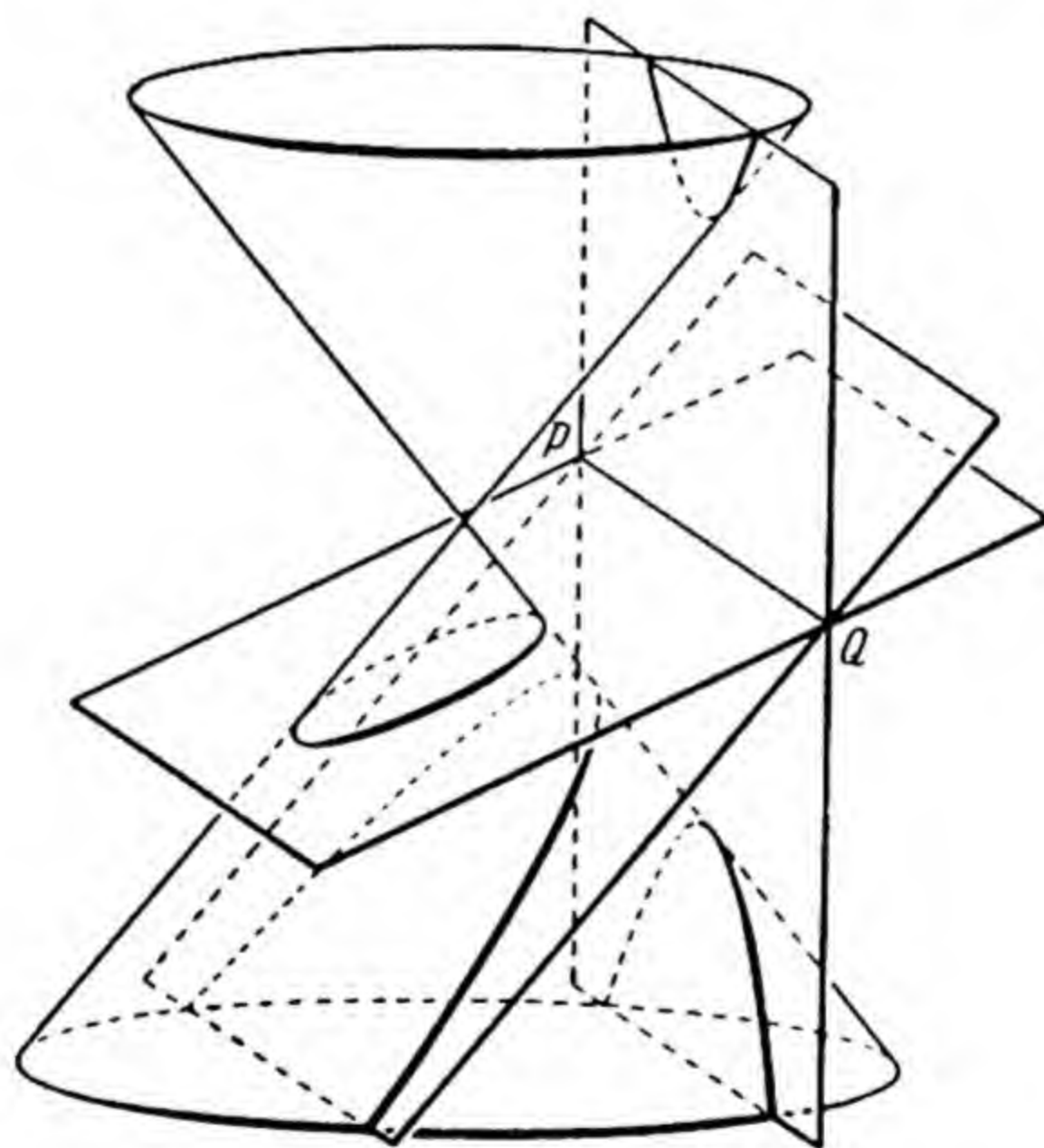


Fig. 71.

This theorem follows from the more general statement that a plane section of a quadric surface is a curve of the second order.

From Fig. 71 it is apparent that, by rotating the cutting plane about the line PQ , the curve of intersection can be made to change. Let it, for example, be an ellipse initially; as the plane rotates, the curve becomes a parabola for the instant when the cutting plane is parallel to a plane tangent to the cone, and then the curve changes to a hyperbola.

Accordingly, ellipses, hyperbolas and parabolas are called *conic sections*,

TRANSFORMATION OF EQUATIONS BY CHANGE OF COORDINATES

§ 41. Examples of Reducing the General Equation of a Second-order Curve to Canonical Form

113. The analysis of the general equation of a second-order curve and its reduction to the simplest (canonical) form constitutes an important problem of analytic geometry. Without attempting to give here a general solution, we shall devote the present section to elucidating the essence of the problem by means of concrete examples.

But first we must make a remark concerning the notation. The general equation of a curve of the second order, that is, the general equation of the second degree in x and y was earlier (§ 15) written as

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

In the theory of second-order curves, however, the majority of formulas contain the coefficients B , D and E divided by 2. It is therefore advisable to put the general equation of the second degree in the form

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0, \quad (1)$$

that is, to denote by the letters B , D and E the halves of the respective coefficients. If, for instance,

$$x^2 + 3xy + 2y^2 + 5x + 4y + 1 = 0$$

is the given equation, then

$$A = 1, \quad B = \frac{3}{2}, \quad C = 2, \quad D = \frac{5}{2}, \quad E = 2, \quad F = 1.$$

The numbers A , B , C , D , E , F are called *the coefficients* of equation (1) (as we see, the application of this term to B , D and E is here a matter of convention). The first three terms of equation (1), that is, the second-degree terms, are referred to as *the highest terms* of the equation.

To illustrate at once the convenience of writing the equation of the second degree in the form (1), note the following identity:

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = \\ = (Ax + By + D)x + (Bx + Cy + E)y + (Dx + Ey + F), \quad (2)$$

which can readily be verified. This identity shows that it is natural to regard the second, fourth and fifth terms of equation (1) as made up of two identical parts each. Relation (2) is helpful in many cases and will presently be used.

114. Let there be given an equation of a second-order curve in the general form

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \quad (1)$$

It is required to simplify this equation by changing the coordinates (that is, by moving the axes to a new and more advantageous position). More precisely, it is required to:

(1) eliminate the term in xy from the equation; (2) reduce the number of the first-degree terms to a minimum (remove them completely, if possible); and (3) remove, if possible, also the constant term. An equation fulfilling these requirements is called *the canonical equation*. Given below are practical illustrations of how to perform the operations necessary for reducing a given equation to its canonical form.

Example. Reduce the equation

$$17x^2 + 12xy + 8y^2 - 46x - 28y + 17 = 0 \quad (3)$$

to its canonical form.

Solution. To begin with, we shall try to simplify the equation by a translation of the coordinate axes. Let us move the origin to the point $S(x_0, y_0)$, which will be regarded as arbitrary for the present. The corresponding transformation of coordinates will be, by § 8,

$$x = \tilde{x} + x_0, \quad y = \tilde{y} + y_0. \quad (4)$$

Introducing the new coordinates in the left-hand member of (3)—i. e., replacing x and y by their expressions (4)—and collecting like terms, we find

$$17x^2 + 12xy + 8y^2 - 46x - 28y + 17 = \\ = 17\tilde{x}^2 + 12\tilde{x}\tilde{y} + 8\tilde{y}^2 + 2(17x_0 + 6y_0 - 23)\tilde{x} + \\ + 2(6x_0 + 8y_0 - 14)\tilde{y} + (17x_0^2 + 12x_0y_0 + 8y_0^2 - \\ - 46x_0 - 28y_0 + 17). \quad (5)$$

The transformed equation of the given curve will be free of first-degree terms if x_0, y_0 are chosen so that the equations

$$17x_0 + 6y_0 - 23 = 0, \\ 6x_0 + 8y_0 - 14 = 0 \quad (6)$$

will hold. Solving these equations simultaneously, we get $x_0 = 1$, $y_0 = 1$. Let \tilde{F} denote the constant term of the transformed equation; computation of \tilde{F} is greatly facilitated by the use of identity (2) together with equations (6):

$$\begin{aligned}\tilde{F} &= 17x_0^2 + 12x_0y_0 + 8y_0^2 - 46x_0 - 28y_0 + 17 = \\ &= (17x_0 + 6y_0 - 23)x_0 + (6x_0 + 8y_0 - 14)y_0 + \\ &+ (-23x_0 - 14y_0 + 17) = -23x_0 - 14y_0 + 17 = -20.\end{aligned}$$

The origin of the new coordinate system is located at the point S (whose old coordinates are $x_0 = 1$, $y_0 = 1$). In terms of the new coordinates, the equation takes the form

$$17\tilde{x}^2 + 12\tilde{x}\tilde{y} + 8\tilde{y}^2 - 20 = 0. \quad (7)$$

Note that the left-hand member of (7) remains unchanged when replacing \tilde{x} , \tilde{y} by $-\tilde{x}$, $-\tilde{y}$. If, therefore, equation (7) is satisfied by some numbers \tilde{x} , \tilde{y} , then it is also satisfied by the numbers $-\tilde{x}$, $-\tilde{y}$. Hence, if a point $M(\tilde{x}, \tilde{y})$ lies on the given curve, then the point $N(-\tilde{x}, -\tilde{y})$ also lies on the curve. But the points $M(\tilde{x}, \tilde{y})$ and $N(-\tilde{x}, -\tilde{y})$ are symmetric with respect to the point S .

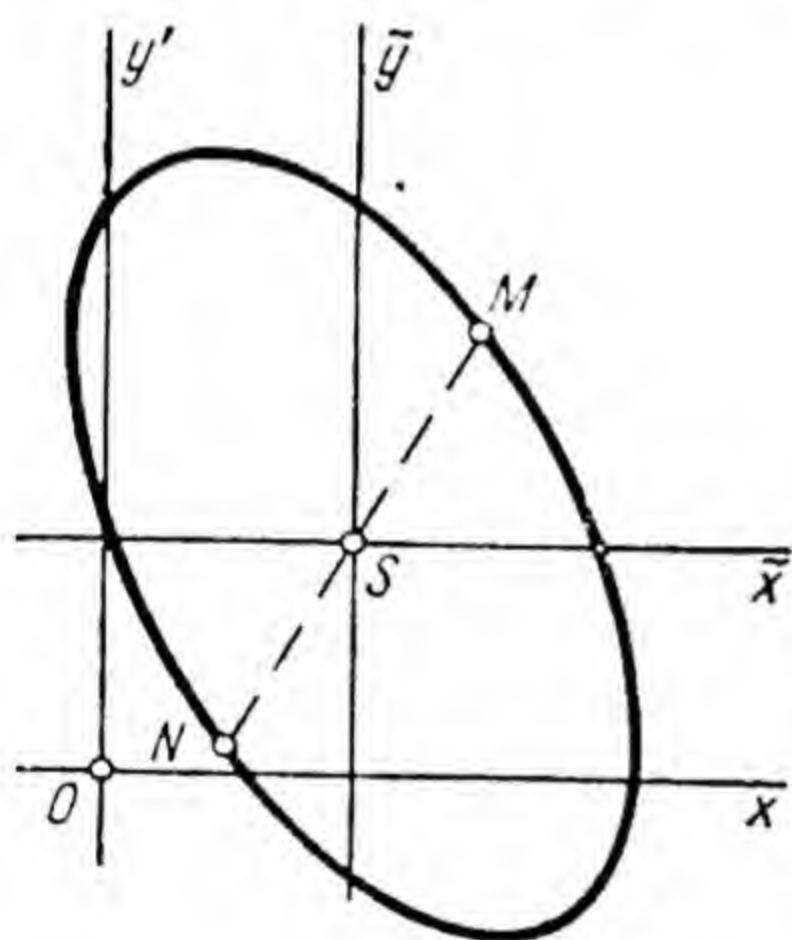


Fig. 72.

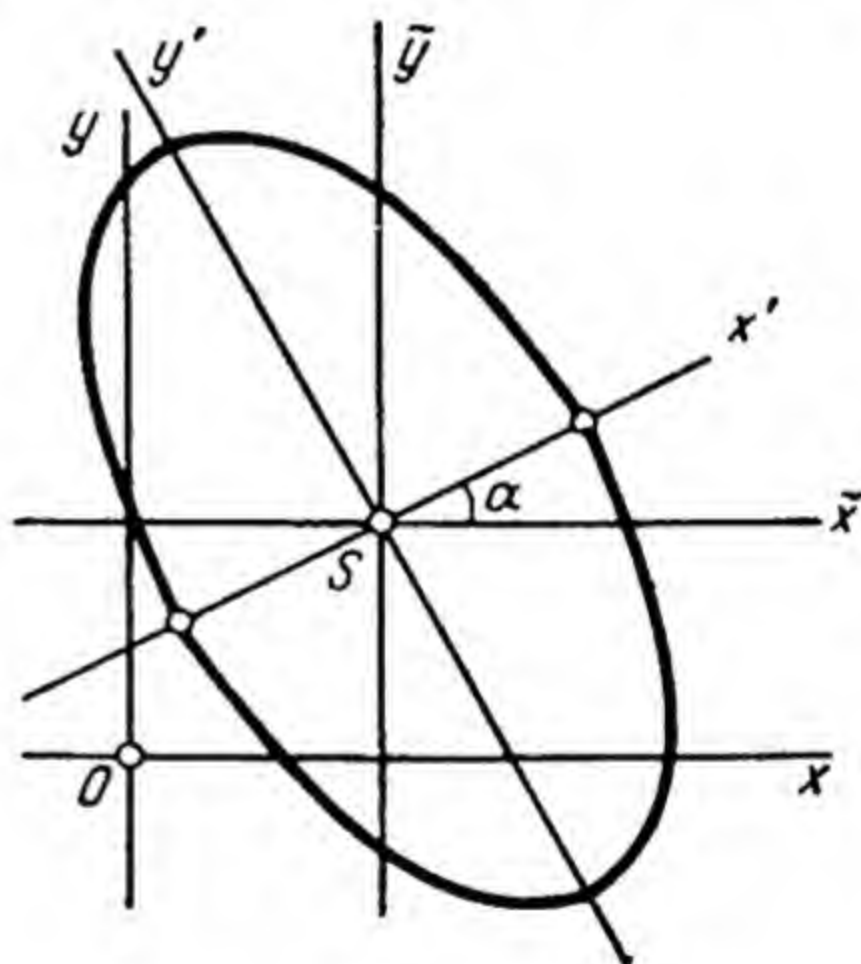


Fig. 73.

Thus, all points of the given curve form pairs symmetric with respect to S (Fig. 72). In this case, the point S is called the centre of symmetry, or simply the centre, of the given curve. The geometric meaning of the transformation performed is now clear: we have moved the origin to the centre of the curve.

Next, we shall rotate the translated axes through a certain angle α . The corresponding transformation of coordinates will be, by § 9,

$$\begin{aligned}\tilde{x} &= x' \cos \alpha - y' \sin \alpha, \\ \tilde{y} &= x' \sin \alpha + y' \cos \alpha.\end{aligned} \quad (8)$$

Replacing \tilde{x} , \tilde{y} in (7) by their expressions (8) and collecting like terms, we obtain

$$\begin{aligned} 17\tilde{x}^2 + 12\tilde{x}\tilde{y} + 8\tilde{y}^2 - 20 = & (17\cos^2\alpha + 12\cos\alpha\sin\alpha + \\ & + 8\sin^2\alpha)x'^2 + 2(-17\cos\alpha\sin\alpha + 6\cos^2\alpha - 6\sin^2\alpha + \\ & + 8\cos\alpha\sin\alpha)x'y' + (17\sin^2\alpha - 12\cos\alpha\sin\alpha + 8\cos^2\alpha)y'^2 - 20. \end{aligned} \quad (9)$$

Let us choose the angle α so that the coefficient of the term in $x'y'$ will vanish. For this purpose, we must solve the trigonometric equation

$$-17\cos\alpha\sin\alpha + 6\cos^2\alpha - 6\sin^2\alpha + 8\cos\alpha\sin\alpha = 0,$$

or

$$6\sin^2\alpha + 9\sin\alpha\cos\alpha - 6\cos^2\alpha = 0.$$

Hence

$$6\tan^2\alpha + 9\tan\alpha - 6 = 0.$$

Solving this last quadratic for $\tan\alpha$, we find: $\tan\alpha = \frac{1}{2}$ or $\tan\alpha = -2$. We shall take the first solution, which corresponds to a rotation of the coordinate axes through an acute angle. Computation of $\cos\alpha$ and $\sin\alpha$ for this value of $\tan\alpha$ gives

$$\begin{aligned} \cos\alpha &= \frac{1}{\sqrt{1+\tan^2\alpha}} = \frac{2}{\sqrt{5}}, \\ \sin\alpha &= \frac{\tan\alpha}{\sqrt{1+\tan^2\alpha}} = \frac{1}{\sqrt{5}}. \end{aligned}$$

Hence, from (9), we find the equation of the given curve in the x' , y' -system:

$$20x'^2 + 5y'^2 - 20 = 0,$$

or

$$\frac{x'^2}{1} + \frac{y'^2}{4} = 1.$$

We have obtained the canonical equation of an ellipse with semi-axes 2 and 1 (the major axis of this ellipse lies on the axis Oy' ; see Fig. 73).

115. In the case of the general curve of the second order

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0,$$

the equations determining its centre $S(x_0, y_0)$ are written as

$$\begin{aligned} Ax_0 + By_0 + D &= 0, \\ Bx_0 + Cy_0 + E &= 0. \end{aligned} \quad (10)$$

After moving the origin to the centre S , the equation of the given curve takes the form

$$A\tilde{x}^2 + 2B\tilde{x}\tilde{y} + C\tilde{y}^2 + \tilde{F} = 0, \quad (11)$$

where

$$\tilde{F} = Ax_0^2 + 2Bx_0y_0 + Cy_0^2 + 2Dx_0 + 2Ey_0 + F.$$

Making use of identity (2), we rewrite this as

$$\tilde{F} = (Ax_0 + By_0 + D)x_0 + (Bx_0 + Cy_0 + E)y_0 + (Dx_0 + Ey_0 + F).$$

Under the condition that x_0, y_0 are the coordinates of the centre of the curve, we find, by (10),

$$\tilde{F} = Dx_0 + Ey_0 + F.$$

To derive equations (10) and (11), the reader should perform (in general form) the operations used in the preceding example to obtain equations (6) and (7).

116. The system of equations (10) may happen to be inconsistent, that is, to have no solutions. In this case, the curve has no centre and the simplification of the given equation must be carried out according to a different plan.

Example. Reduce the equation

$$4x^2 - 4xy + y^2 - 2x - 14y + 7 = 0 \quad (12)$$

to its canonical form.

Solution. On forming equations (10):

$$\begin{aligned} 4x_0 - 2y_0 - 1 &= 0, \\ -2x_0 + y_0 - 7 &= 0, \end{aligned}$$

we see that the obtained system is inconsistent. Hence the given curve has no centre, and so we cannot proceed as in Art. 114.

We shall use a different procedure. Prior to changing the origin, let us rotate the axes through a certain angle α . By § 9, the corresponding formulas of coordinate transformation will be

$$\begin{aligned} x &= x' \cos \alpha - y' \sin \alpha, \\ y &= x' \sin \alpha + y' \cos \alpha. \end{aligned}$$

Introducing the new coordinates in the left-hand member of (12), we get

$$\begin{aligned} 4x^2 - 4xy + y^2 - 2x - 14y + 7 &= (4 \cos^2 \alpha - 4 \cos \alpha \sin \alpha + \\ &+ \sin^2 \alpha) x'^2 + 2(-4 \sin \alpha \cos \alpha - 2 \cos^2 \alpha + 2 \sin^2 \alpha + \\ &+ \sin \alpha \cos \alpha) x'y' + (4 \sin^2 \alpha + 4 \sin \alpha \cos \alpha + \cos^2 \alpha) y'^2 + \\ &+ 2(-\cos \alpha - 7 \sin \alpha) x' + 2(\sin \alpha - 7 \cos \alpha) y' + 7. \end{aligned} \quad (13)$$

Let us now choose the angle α so that the coefficient of $x'y'$ will vanish. To this end, we must solve the trigonometric equation

$$-4 \sin \alpha \cos \alpha - 2 \cos^2 \alpha + 2 \sin^2 \alpha + \sin \alpha \cos \alpha = 0,$$

or

$$2 \sin^2 \alpha - 3 \sin \alpha \cos \alpha - 2 \cos^2 \alpha = 0.$$

Hence

$$2 \tan^2 \alpha - 3 \tan \alpha - 2 = 0,$$

This gives $\tan \alpha = 2$ and $\tan \alpha = -\frac{1}{2}$. We take the first solution, which corresponds to a rotation of the axes through an acute angle. Calculating $\cos \alpha$ and $\sin \alpha$ for this value of $\tan \alpha$, we get

$$\cos \alpha = \frac{1}{\sqrt{1 + \tan^2 \alpha}} = \frac{1}{\sqrt{5}},$$

$$\sin \alpha = \frac{\tan \alpha}{\sqrt{1 + \tan^2 \alpha}} = \frac{2}{\sqrt{5}}.$$

Hence, from (13), we find the equation of the given curve in the x', y' -system:

$$5y'^2 - 6\sqrt{5}x' - 2\sqrt{5}y' + 7 = 0. \quad (14)$$

To effect a further simplification of (14), we shall now translate the axes Ox', Oy' .

Rewrite (14) as follows:

$$5\left(y'^2 - 2\frac{\sqrt{5}}{5}y'\right) - 6\sqrt{5}x' + 7 = 0.$$

Completing the square in y' , we obtain

$$\left(y' - \frac{\sqrt{5}}{5}\right)^2 - \frac{6\sqrt{5}}{5}\left(x' - \frac{\sqrt{5}}{5}\right) = 0.$$

Let us once again introduce new coordinates (x'', y'') , setting

$$x' = x'' + \frac{\sqrt{5}}{5}, \quad y' = y'' + \frac{\sqrt{5}}{5},$$

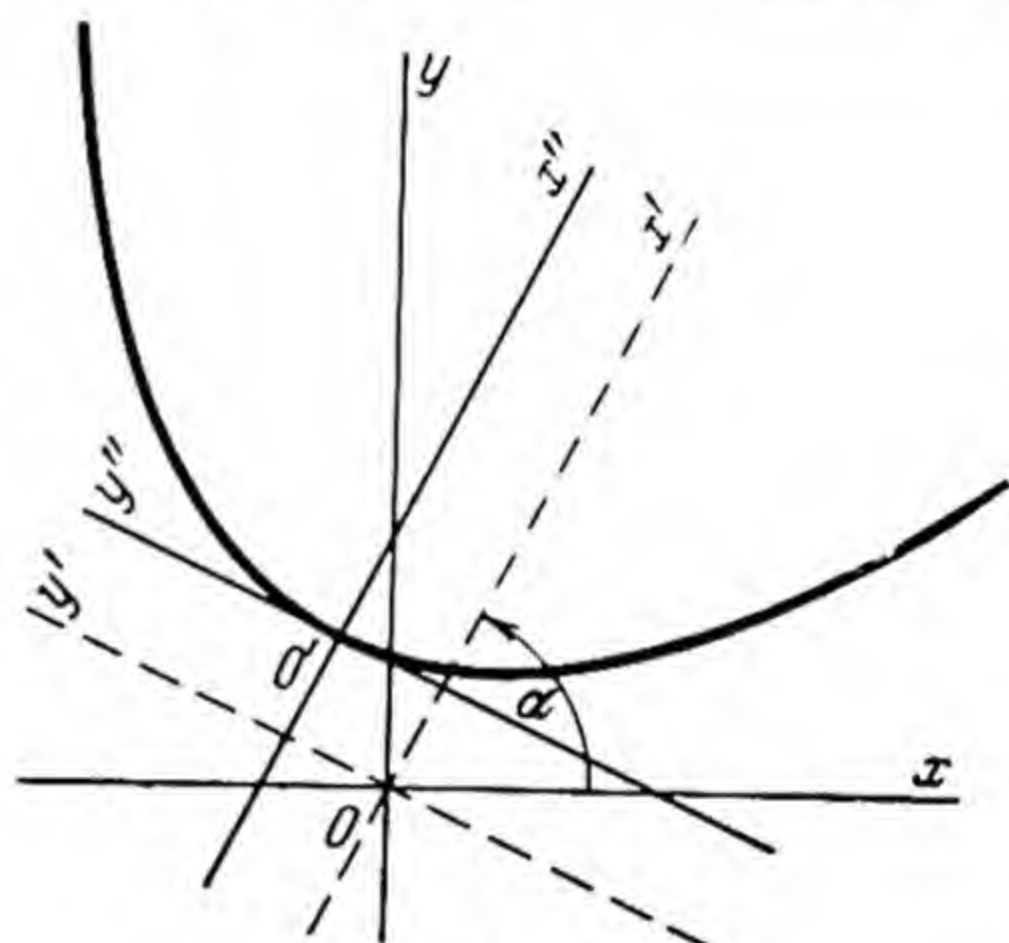


Fig. 74.

which corresponds to translating the axes by the amount $\frac{\sqrt{5}}{5}$ in the direction of the axis Ox' , and by the same amount in the direction of the axis Oy' . In terms of x'', y'' , the equation of the given curve will take the form

$$y''^2 = \frac{6\sqrt{5}}{5}x''.$$

This is the canonical equation of the parabola with parameter $p = \frac{3\sqrt{5}}{5}$ and with vertex at the origin of the x'', y'' -coordinate system; the parabola is symmetric with respect to the x'' -axis and opens in the positive direction of this axis. The coordinates of the vertex are $\left(\frac{\sqrt{5}}{5}, \frac{\sqrt{5}}{5}\right)$ in the x', y' -system, and $\left(-\frac{1}{5}, \frac{3}{5}\right)$ in the x, y -system. The location of the parabola is shown in Fig. 74.

117. Let us return to the system of equations (10) representing the centre of the given curve:

$$\begin{aligned} Ax_0 + By_0 + D &= 0, \\ Bx_0 + Cy_0 + E &= 0. \end{aligned} \quad (10)$$

Denote by δ the determinant of this system:

$$\delta = \begin{vmatrix} A & B \\ B & C \end{vmatrix} = AC - B^2.$$

If $\delta \neq 0$, the system (10) has a unique solution (see Appendix, § 1). In this case, the given second-order curve has a single centre and is called a *central curve*. Central curves include ellipses and hyperbolas. When $\delta \neq 0$, it may, however, also happen that the canonical form, to which the given equation is reducible, resembles that of an ellipse or a hyperbola, but is fully identical with neither, as in the three examples below. Before considering them, it should be noted that, if $\delta \neq 0$, the general equation of the second degree can always be simplified by following the same procedure as in the example of Art. 114. The transformation work is therefore left out in the examples that follow.

Example 1. The equation $5x^2 + 6xy + 5y^2 - 4x + 4y + 12 = 0$ (for which $\delta = 9 \neq 0$) reduces to the canonical form $x'^2 + 4y'^2 + 4 = 0$, or

$$\frac{x'^2}{4} + \frac{y'^2}{1} = -1.$$

This equation bears a resemblance to the canonical equation of an ellipse. However, it represents no real geometric figure in the plane since, for any real numbers x' , y' , its left-hand member is non-negative, whereas its right-hand member is -1 . This and similar equations are referred to as the equations of an *imaginary ellipse*.

Example 2. The equation $5x^2 + 6xy + 5y^2 - 4x + 4y + 4 = 0$ (for which $\delta = 9 \neq 0$) reduces to the canonical form $x'^2 + 4y'^2 = 0$, or

$$\frac{x'^2}{4} + \frac{y'^2}{1} = 0. \quad (*)$$

Equation (*) also bears a resemblance to the canonical equation of an ellipse, but it represents a single point ($x' = 0$, $y' = 0$), and not an ellipse. This and similar equations are called the equations of a *degenerate ellipse*. To see what is meant by this term, consider the equation

$$\frac{x'^2}{4} + \frac{y'^2}{1} = \varepsilon^2, \quad (**)$$

where ε is any positive number ($\varepsilon > 0$). Equation (**) represents an ordinary ellipse with semi-axes $a = 2\varepsilon$, $b = \varepsilon$. Now imagine that ε tends to zero.

Then $a \rightarrow 0$, $b \rightarrow 0$, and the ellipse "degenerates" into a point (Fig. 75), while equation (**) is transformed into equation (*).

Example 3. The equation $3x^2 + 10xy + 3y^2 + 16x + 16y + 16 = 0$ (for which $\delta = -16 \neq 0$) reduces to the canonical form $x'^2 - 4y'^2 = 0$, or

$$\frac{x'^2}{4} - \frac{y'^2}{1} = 0. \quad (*)$$

Equation (*), which resembles the canonical equation of a hyperbola, represents a pair of intersecting straight lines: $x' - 2y' = 0$, $x' + 2y' = 0$. This and similar equations are called the equations of a *degenerate hyperbola*.

To see what is meant by the term, let

$$\frac{x'^2}{4} - \frac{y'^2}{1} = \epsilon^2, \quad (**)$$

where ϵ is any positive number ($\epsilon > 0$). Equation (**) represents an ordinary hyperbola with semi-axes $a = 2\epsilon$, $b = \epsilon$, and with vertices on the x -axis. Now imagine that ϵ tends to zero. Then $a \rightarrow 0$, $b \rightarrow 0$, the vertices of the hyperbola

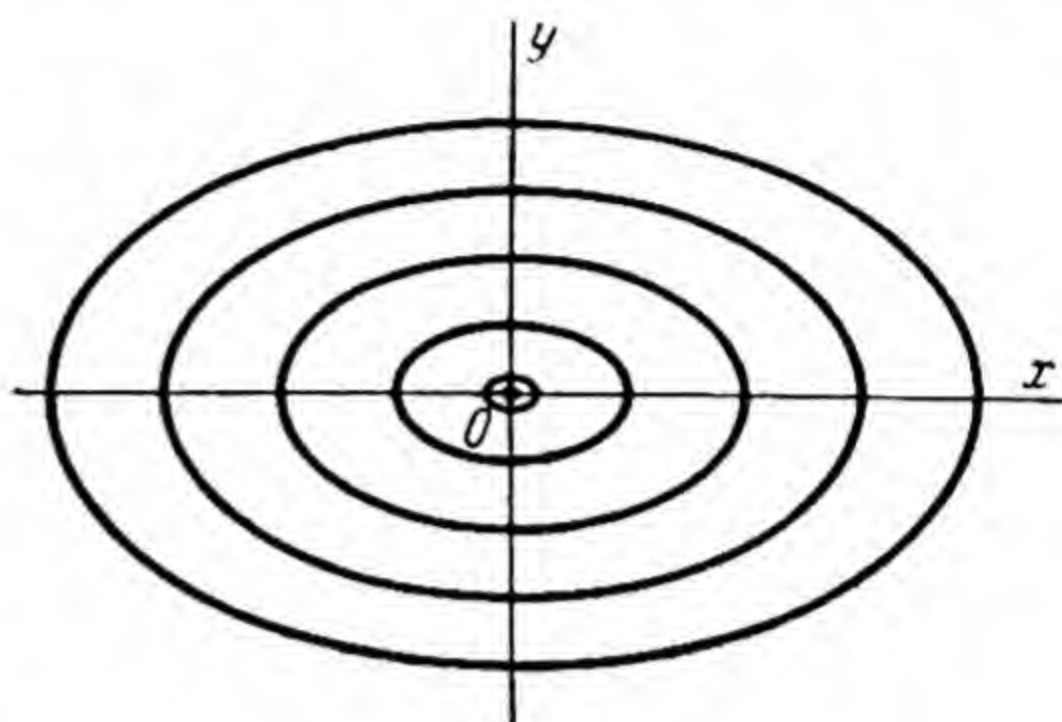


Fig. 75.

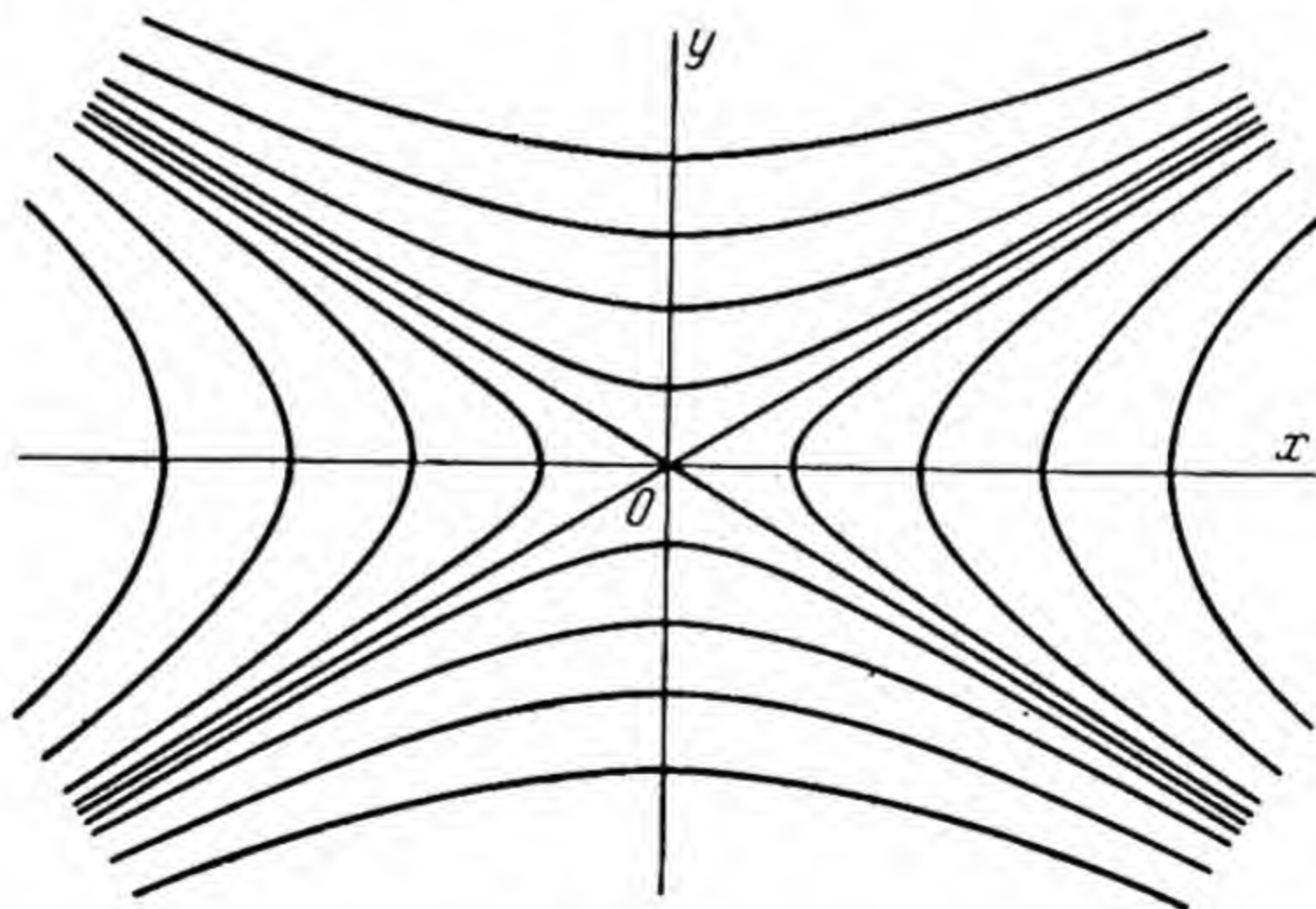


Fig. 76.

come closer and closer together, and the hyperbola "degenerates" into a pair of lines, namely, into the pair of its asymptotes, while equation (**) is transformed into equation (*). If ϵ^2 is replaced in (**) by $-\epsilon^2$, we obtain a hyperbola with vertices on the y -axis. As $\epsilon \rightarrow 0$, this hyperbola degenerates into the same pair of lines as before (Fig. 76).

Suppose now that, for the given general equation of the second degree, we have $\delta = 0$. If $\delta = 0$, the following two cases are possible:

(1) The system of equations (10) has no solutions at all; then our second-order curve has no centre. In this case, the given equation can always be reduced to its canonical form by proceeding as shown in the example of Art. 116, the result being always the canonical equation of a parabola.

(2) The system of equations (10) has an infinite number of solutions; then the given second-order curve has an infinite number of centres.

Example 4. Consider the second-order curve

$$4x^2 - 4xy + y^2 + 4x - 2y - 3 = 0, \quad (*)$$

for which $\delta = 0$. In this case, the system (10) will be

$$\begin{aligned} 4x_0 - 2y_0 + 2 &= 0, \\ -2x_0 + y_0 - 1 &= 0. \end{aligned}$$

This system is equivalent to a single equation, $2x_0 - y_0 + 1 = 0$; consequently, the curve has an infinite number of centres, which form the straight line $2x - y + 1 = 0$. Note that the left-hand member of equation (*) is factorable into two first-degree expressions:

$$\begin{aligned} 4x^2 - 4xy + y^2 + 4x - 2y - 3 &= \\ &= (2x - y + 3)(2x - y - 1). \end{aligned}$$

Hence, the curve under consideration is a pair of parallel straight lines,

$$2x - y + 3 = 0 \text{ and } 2x - y - 1 = 0.$$

The straight line $2x - y + 1 = 0$, made up of the centres, is simply the mean line of this pair of lines (Fig. 77).

To simplify the given equation (*), we can proceed as in Art. 116. Transforming

the left-hand member of the equation in a manner analogous to that used in (13) and applying further the same argument and procedure, we find $\tan \alpha = 2$. By rotating the axes through the angle α ($\tan \alpha = 2$), we reduce the given equation to the form

$$5y'^2 - 2\sqrt{5}y' - 3 = 0;$$

hence

$$5\left(y' - \frac{\sqrt{5}}{5}\right)^2 - 4 = 0.$$

Setting $x' = x''$, $y' = y'' + \frac{\sqrt{5}}{5}$, which corresponds to a translation of the axes Ox' , Oy' by the amount $\frac{\sqrt{5}}{5}$ in the direction of the axis Oy' , we

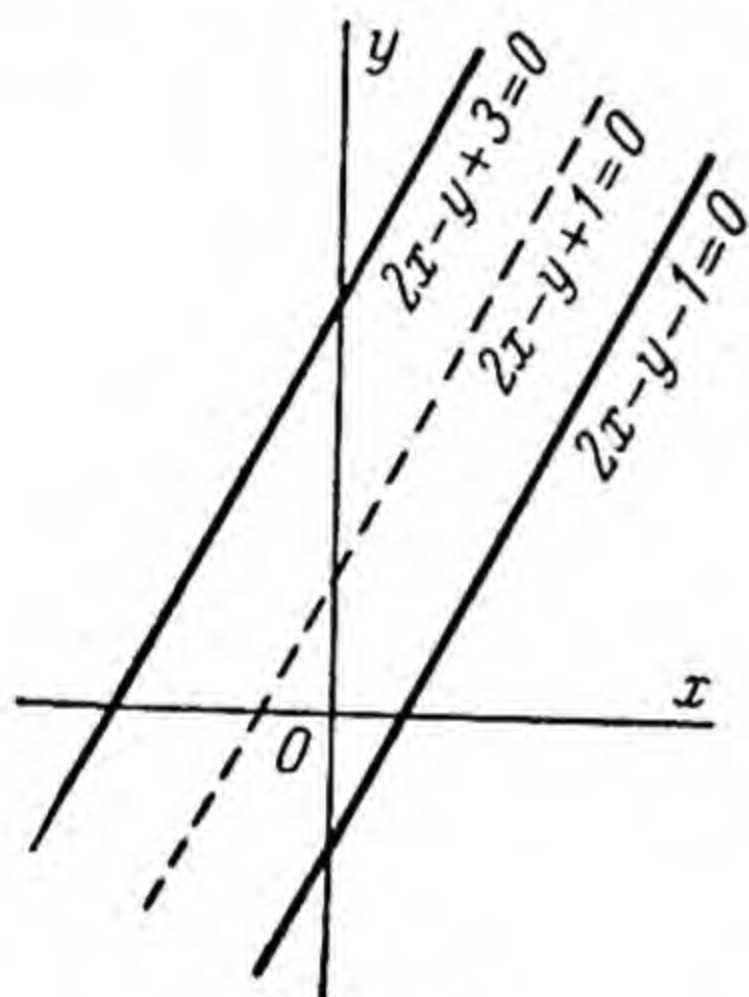


Fig. 77.

finally get

$$5y''^2 - 4 = 0.$$

Once again, the given equation represents a pair of parallel lines ($\sqrt{5}y'' - 2 = 0$ and $\sqrt{5}y'' + 2 = 0$, when referred to the x'', y'' -coordinate system).

An equation of the second degree, which represents a second-order curve with an infinite number of centres (as in the last example), is customarily referred to as the equation of a *degenerate parabola*.

118. The examples considered in this section show convincingly enough that the general equation of a second-order curve can always be reduced to the canonical form.

§ 42. The Hyperbola as the Inverse Proportionality Graph. The Parabola as the Graph of a Quadratic Function

119. In mathematics and its applications, one often encounters an equation of the form $xy = m$, or $y = \frac{m}{x}$ (where $m = \text{const} \neq 0$); which is referred to as *the equation expressing the inverse proportionality of the quantities x and y* . It is easy to show that, in rectangular cartesian coordinates x, y , such an equation represents an equilateral hyperbola whose asymptotes coincide with the coordinate axes.

For, if we rotate the axes Ox and Oy through an angle $\alpha = 45^\circ$, the coordinates of all points in the plane will be transformed according to the formulas

$$\left. \begin{aligned} x &= x' \cos \alpha - y' \sin \alpha = \frac{x' - y'}{\sqrt{2}}, \\ y &= x' \sin \alpha + y' \cos \alpha = \frac{x' + y'}{\sqrt{2}}. \end{aligned} \right\} \quad (1)$$

Transforming the equation $xy = m$ by formulas (1), we obtain, in terms of the new coordinates,

$$\frac{x'^2}{2m} - \frac{y'^2}{2m} = 1.$$

We see that this is the canonical equation of an equilateral hyperbola with semi-axes $a = b = \sqrt{2|m|}$; its asymptotes are inclined at 45° to the new coordinate axes and, consequently, are coincident with the original axes; if the number m is positive, then our hyperbola cuts the new x' -axis; if m is negative, the hyperbola

cuts the new y' -axis. Hence we conclude that, as was to be shown, the equation $xy = m$ represents an equilateral hyperbola whose asymptotes coincide with the coordinate axes; this hyperbola is situated in the first and third quadrants if $m > 0$ (Fig. 78a), or in the second and fourth quadrants if $m < 0$ (Fig. 78b).

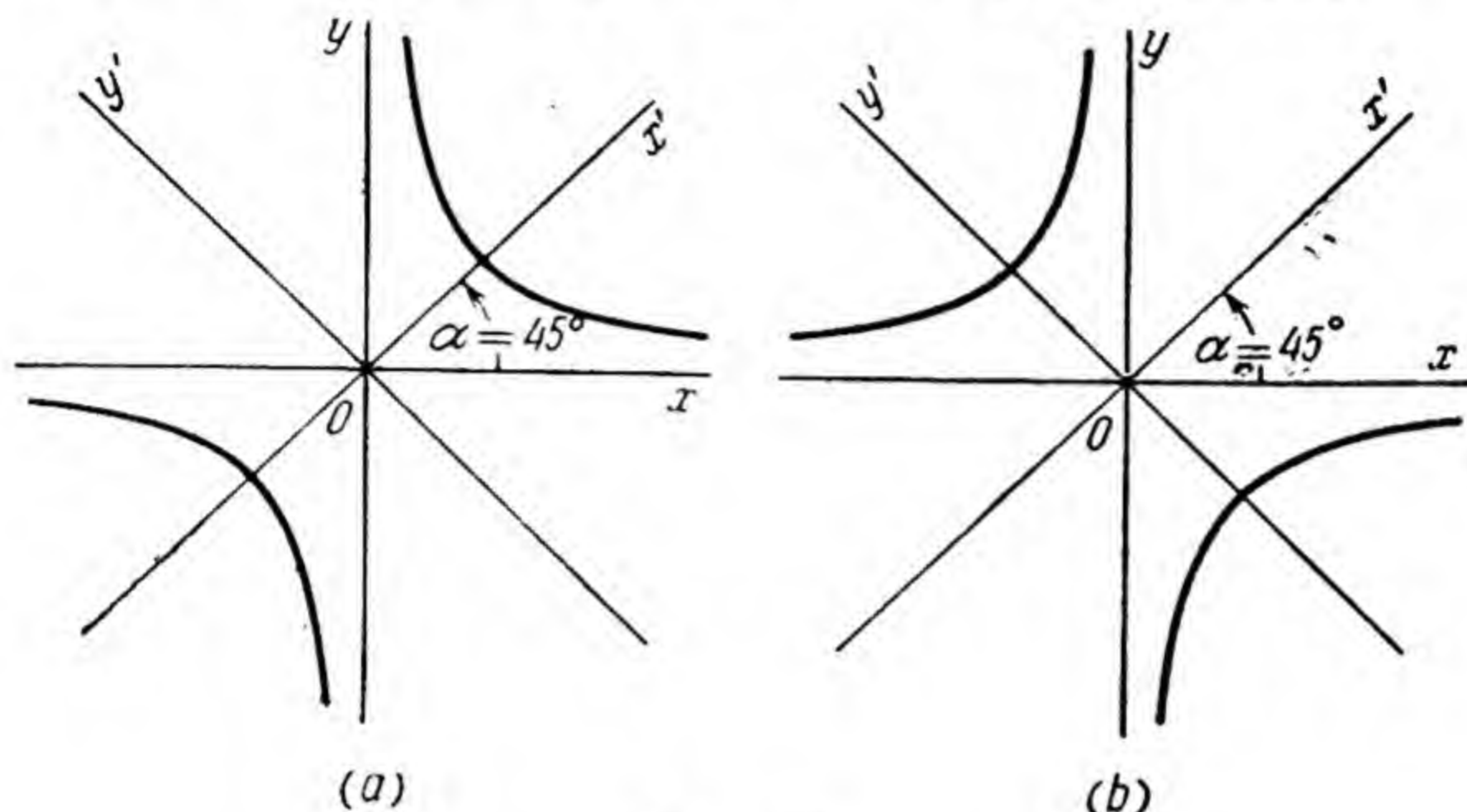


Fig. 78.

Accordingly, an *equilateral hyperbola* may also be spoken of as *the inverse proportionality graph*.

120. The equation

$$y = ax^2 + bx + c, \quad (2)$$

where $a \neq 0$, represents a parabola whose axis of symmetry is perpendicular to the x -axis; if $a > 0$, the parabola opens upwards; if $a < 0$, the parabola opens downwards.

To prove this, we have merely to reduce equation (2) to its canonical form. For this purpose, rewrite the equation as

$$y = a \left(x^2 + \frac{b}{a} x \right) + c,$$

or

$$y = a \left(x^2 + \frac{b}{a} x + \frac{b^2}{4a^2} \right) + c - \frac{b^2}{4a}, \quad (3)$$

or

$$y - \frac{4ac - b^2}{4a} = a \left(x + \frac{b}{2a} \right)^2.$$

Now let the origin be moved to the point $\left(-\frac{b}{2a}, \frac{4ac - b^2}{4a} \right)$.

Then the coordinates of all points in the plane will be transformed according to the formulas

$$x = \tilde{x} - \frac{b}{2a}, \quad y = \tilde{y} + \frac{4ac - b^2}{4a},$$

and equation (3), in the new coordinates, will take the form

$$\tilde{y} = a\tilde{x}^2,$$

or

$$\tilde{x}^2 = \pm 2p\tilde{y}, \quad (4)$$

where p is a positive number determined by the relation $\pm p = \frac{1}{2a}$.

We have obtained the canonical equation of a parabola with vertex at the new origin. This parabola is symmetric with respect to the new axis of ordinates (the \tilde{y} -axis) and opens upwards or downwards according as the number $a = \frac{1}{\pm 2p}$ is positive or negative. Since the \tilde{y} -axis is perpendicular to the original x -axis, the parabola is situated exactly as was indicated at the beginning of the article. Our assertion is thus proved.

121. The expression $ax^2 + bx + c$ is called a *quadratic function* of x . Accordingly, a *parabola* (with vertical axis) can be referred to as the *graph of a quadratic function*.

Example. The equation $y = 2x^2 - 4x - 1$ represents a parabola which opens upwards, since $a = 2 > 0$. In order to find its vertex, rewrite the equation as $y + 3 = 2(x - 1)^2$. To reduce this equation to its canonical form, the origin must be moved to the point $(1, -3)$. This point is the vertex of our parabola (Fig. 79).

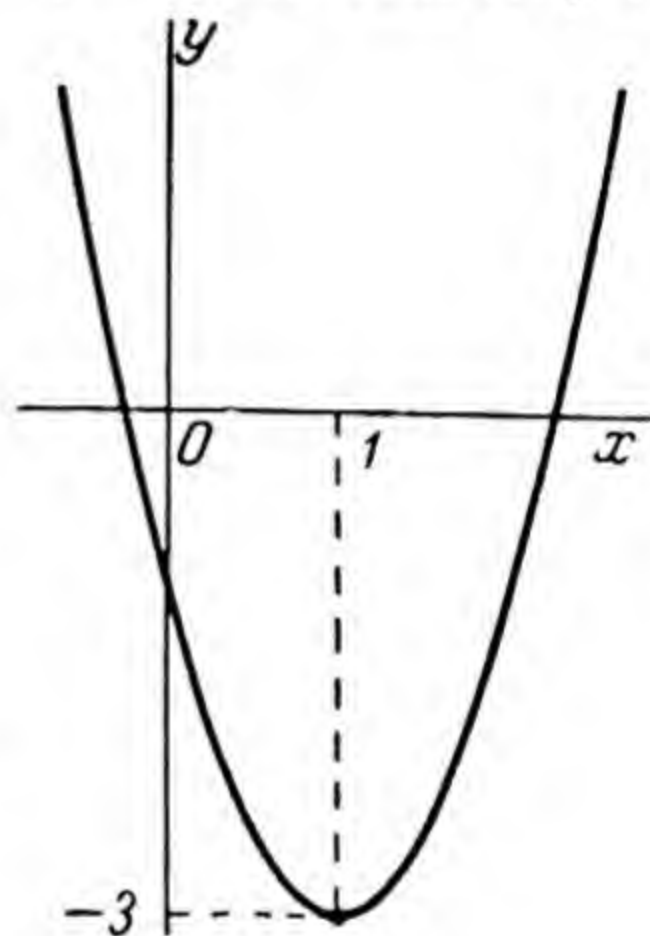


Fig. 79.

PART TWO

**Solid Analytic
Geometry**

SOME ELEMENTARY PROBLEMS OF SOLID ANALYTIC GEOMETRY

§ 43. Rectangular Cartesian Coordinates in Space

122. When a method has been indicated, which permits us to determine the location of points in space by the specification of numbers, then we say that a coordinate system has been attached to space. We shall now consider the simplest and most commonly used coordinate system, called the rectangular cartesian system of coordinates.

A rectangular cartesian coordinate system in space is determined by the choice of a linear unit (for measurement of lengths) and of three concurrent and mutually perpendicular axes, numbered (that is, designated as the first, second and third axis) in any order.

The point of intersection of the axes is called *the origin of coordinates*, and the axes themselves are called *the coordinate axes*; the first coordinate axis is also termed the *x-axis* or *axis of abscissas*, the second is termed the *y-axis* or *axis of ordinates*, and the third is termed the *z-axis* or *axis of applicates*.

Let us denote the origin by the letter O , the x -axis by the letters Ox , the y -axis by Oy , and the z -axis by Oz . In diagrams, the letters x , y , z mark the respective axes at the points farthest from O in the positive direction, so that the direction of each axis is unambiguously indicated by the position of the letters.

Let M be an arbitrary point in space. Project the point M on the coordinate axes, that is, drop perpendiculars from M to the lines Ox , Oy and Oz ; designate the feet of these perpendiculars as M_x , M_y and M_z , respectively.

The coordinates of a point M in a given system are defined as the numbers

$$x = OM_x, \quad y = OM_y, \quad z = OM_z,$$

where OM_x is the value of the segment $\overline{OM_x}$ of the x -axis, OM_y is the value of the segment $\overline{OM_y}$ of the y -axis, and OM_z is the value of the segment $\overline{OM_z}$ of the z -axis (the definition of the value of an axis segment was given in Art. 2). The number x is called the first coordinate or abscissa of the point M , the number y is

called the second coordinate or ordinate of M , and the number z is called the third coordinate or applicate of M . In the text, the coordinates of a point are given in parentheses, next to the letter denoting the point itself: $M(x, y, z)$.

The projection of the point M on the axis Ox can also be obtained as follows: drop a perpendicular from M to the plane Oxy and, from its foot M_{xy} , draw a perpendicular to the axis Ox ; this second perpendicular will have M_x as its foot; in other words, M_x is the projection of M_{xy} on the axis Ox . The projection of M_{xy} on the axis Oy is, obviously, the point M_y .

Similarly, if M_{xz} and M_{yz} are the respective feet of the perpendiculars dropped from M to the planes Oxz and Oyz , then the points M_x , M_y and M_z will be obtained by projecting M_{xz} and M_{yz} on the respective coordinate axes. (By this method, each of the points M_x , M_y , M_z can be obtained in two ways; for example, the point M_x is the projection on the axis Ox of both M_{xy} and M_{xz} .)

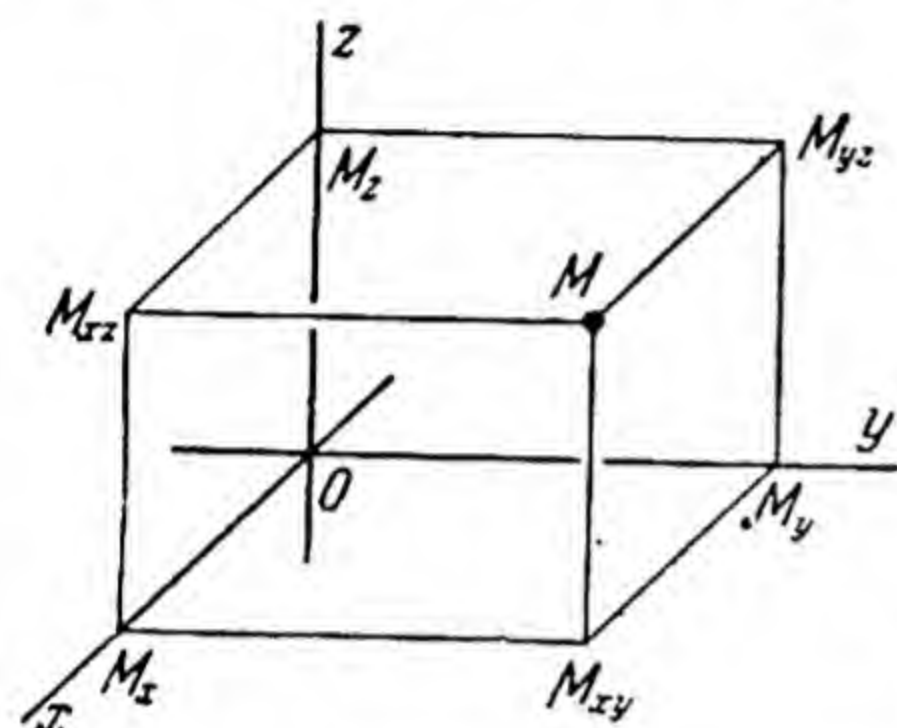


Fig. 80.

The points M_x , M_y , M_z , M_{xy} , M_{xz} , M_{yz} , and O form the vertices of a rectangular parallelepiped, whose edges, taken with the appropriate signs, are the coordinates of the point M . This parallelepiped is shown in Fig. 80.

123. When a rectangular cartesian system of coordinates has been attached to space, then each point of space has one completely determined triad of coordinates x , y , z in this system. Conversely, for any three(real) numbers x , y , z , there will be found in space one completely determined point with the abscissa x , the ordinate y and the applicate z . To plot a point from its coordinates x , y , z , the segment $\overline{OM_x}$, equal in value to x , is laid off (from the origin) on the x -axis, the segment $\overline{OM_y}$ of the value y is laid off on the y -axis, and the segment $\overline{OM_z}$ of the value z is laid off on the z -axis. Then, passing through M_x , M_y and M_z planes perpendicular to the axes Ox , Oy and Oz , respectively, we shall find the required point as the point of intersection of these three planes.

124. Let us now agree about the use of certain terms (assuming that the axes are chosen as in Fig. 80).

The plane Oyz divides all space into two half-spaces; the half-space containing the positive half of the axis Ox will be termed *the near half-space*, and the other half-space will be termed *the far half-space*.

Similarly, the plane Oxz divides space into two half-spaces, of which the one containing the positive half of the axis Oy will be termed *the right half-space*, and the other, *the left half-space*.

Finally, the plane Oxy also divides all space into two half-spaces, of which the one containing the positive half of the axis Oz will be termed *the upper half-space*, and the other, *the lower half-space*.

125. Let M be an arbitrary point of the near half-space; then the segment \overline{OM}_x is positively directed on the axis Ox and, consequently, the abscissa $x = OM_x$ of the point M is positive. If, on the other hand, M is situated in the far half-space, then the segment \overline{OM}_x is negatively directed on the axis Ox and the number $x = OM_x$ is negative. Finally, if the point M lies in the plane Oyz , its projection M_x on the axis Ox coincides with the point O , so that $x = OM_x$ is zero.

Thus, *all points of the near half-space have positive abscissas ($x > 0$); all points of the far half-space have negative abscissas ($x < 0$); the abscissas of all points lying in the plane Oyz are equal to zero ($x = 0$).*

By reasoning in a similar way, it can easily be established that *all points of the right half-space have positive ordinates ($y > 0$); all points of the left half-space have negative ordinates ($y < 0$); the ordinates of all points lying in the plane Oxz are equal to zero ($y = 0$).*

Finally, *all points of the upper half-space have positive applicates ($z > 0$); all points of the lower half-space have negative applicates ($z < 0$); the applicates of all points lying in the plane Oxy are equal to zero ($z = 0$).*

Since the points of the plane Oxz are characterised by the relation $y = 0$, and the points of the plane Oxy , by the relation $z = 0$, we conclude that *the points of Ox are characterised by the two relations*

$$y = 0, \quad z = 0.$$

In like manner, the points of the axis Oy are characterised by the two relations

$$x = 0, \quad z = 0,$$

and the points of the axis Oz by

$$x = 0, \quad y = 0.$$

Note that the origin O , as the point of intersection of the axes, has all the three of its coordinates equal to zero ($x = 0$, $y = 0$, $z = 0$) and is characterised by this property (that is, all three coordinates are zero only for the point O).

126. The three planes Oxy , Oxz and Oyz jointly divide space into eight parts, called *octants* and numbered according to the following rule: The first octant is the one lying simultaneously in the near, right and upper half-spaces; the second octant lies in the far, right and upper half-spaces; the third octant lies in the far, left and upper half-spaces; the fourth octant lies in the near, left and upper half-spaces; the fifth, sixth, seventh and eighth octants are those situated in the lower half-space under the first, second, third and fourth octants, respectively.

Let M be a point with coordinates x , y , z . From the foregoing, it follows that,

- if $x > 0$, $y > 0$, $z > 0$, M lies in the first octant;
- if $x < 0$, $y > 0$, $z > 0$, M lies in the second octant;
- if $x < 0$, $y < 0$, $z > 0$, M lies in the third octant;
- if $x > 0$, $y < 0$, $z > 0$, M lies in the fourth octant;
- if $x > 0$, $y > 0$, $z < 0$, M lies in the fifth octant;
- if $x < 0$, $y > 0$, $z < 0$, M lies in the sixth octant;
- if $x < 0$, $y < 0$, $z < 0$, M lies in the seventh octant;
- if $x > 0$, $y < 0$, $z < 0$, M lies in the eighth octant.

Consideration of the coordinate half-spaces and octants is useful in that it permits an easy orientation as to the position of the given points by the signs of their coordinates.

§ 44. The Concept of a Free Vector. The Projection of a Vector on an Axis

127. From elementary physics the reader knows that some physical quantities, such as temperature, mass, density, are called *scalar* quantities. Some other quantities, such as force, displacement, velocity, acceleration, are called *vector* quantities.

Every *scalar* quantity may be characterised by a single *number*, which represents the ratio of this quantity to an appropriate unit of measure. On the other hand, the specification of a *number* is *not enough* for characterising a *vector* quantity, since vector quantities, apart from being dimensional, are also *directional* quantities.

Geometric vectors serve for abstract expression of concrete (physical) vector quantities.

Geometric vectors, or simply *vectors*, are defined as directed line segments.

Geometric vectors form the subject of the so-called vector analysis in the same way as numbers form the subject of arithmetics. In vector analysis, certain operations are performed on vectors; these operations are mathematical abstractions of certain uniform operations performed on various concrete vector quantities in physics.

Vector analysis, originally developed to satisfy the needs of physics, has also proved to be fruitful in mathematics itself. In this book vectors are used as one of convenient tools of analytic geometry.

Initial information on vector analysis is contained in the next chapter. The remaining articles of the present chapter deal only with the elementary, purely geometric propositions on directed segments in space.

However, it seems advisable already at this stage to introduce some of the notions, designations and terms adopted in vector analysis.

128. Since a vector is a directed segment, it will, as before, be designated in the text by two capital letters with a bar over them, the first of the letters denoting the initial point, and the second letter the terminal point of the vector. In addition, we shall very frequently indicate a vector by a single small letter in bold-face type (as, for example, ***a***). In diagrams, a vector will always be represented by an *arrow*; if the vector is designated in the text by a single letter, this letter will be placed at the head of the arrow. The initial point of a vector is also called its *point of application*. A vector whose initial and terminal points coincide is called a *zero vector*. Vectors lying on the same straight line or on parallel lines are said to be *collinear*.

129. Definition of the equality of vectors. *Vectors are called equal if they are collinear and have the same length and direction.*

The vectors \overline{AB} and \overline{CD} shown in Fig. 81 are equal ($\overline{AB} = \overline{CD}$ *); Fig. 82 presents *unequal* vectors \overline{PQ} and \overline{PR} ($\overline{PQ} \neq \overline{PR}$), \overline{EF} and \overline{GH} ($\overline{EF} \neq \overline{GH}$).

Obviously, *two vectors each equal to a third vector are equal to each other.*

From the definition of the equality of vectors, it follows that, for any vector ***a*** and any point *P* whatsoever, there exists one, and only one, vector \overline{PQ} extending from *P* and equal to the vector ***a***; in other words, *the point of application of every vector may*

*) The vectors are supposed to lie in the plane of the drawing.

be chosen at pleasure. Accordingly, vectors are considered in geometry as determined to within their position (that is, no difference is made between equal vectors obtained from one another by translation). It is in this sense that vectors are said to be *free*.

130. The length of a vector (in a given scale) is called its *modulus*. The modulus of a zero vector is equal to zero. The modulus of a vector \mathbf{a} is designated as $|\mathbf{a}|$ or a . Clearly, if $\mathbf{a} = \mathbf{b}$,

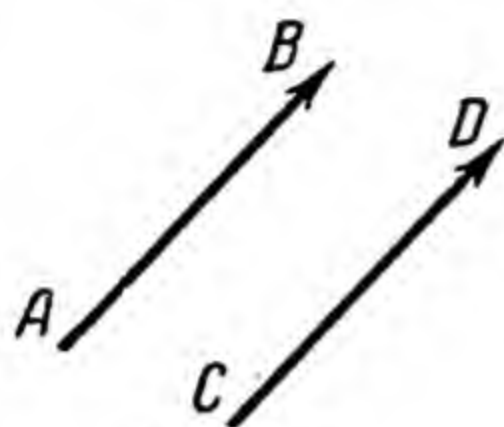


Fig. 81.

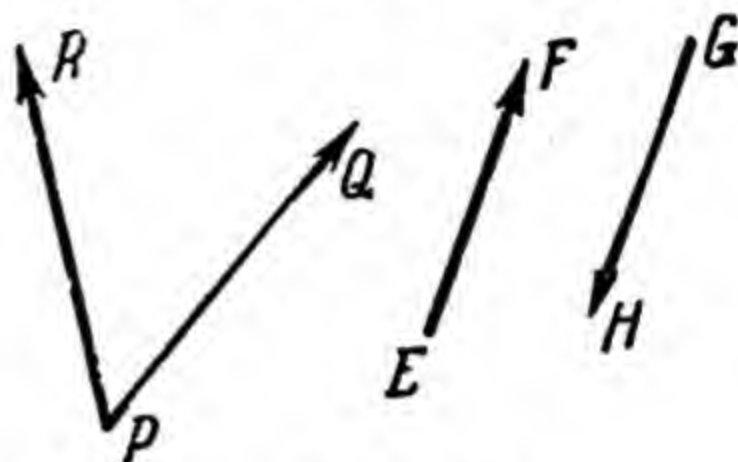


Fig. 82.

then $|\mathbf{a}| = |\mathbf{b}|$; the converse conclusion is, of course, impermissible.

131. Let there be given an arbitrary axis u and a vector \overline{AB} . From the points A and B , drop perpendiculars to the axis u and denote their feet by A' and B' , respectively. The number $A'B'$, that is, the value of the directed segment $\overline{A'B'}$ of the axis u , is the projection of the vector \overline{AB} on the axis u :

$$\text{proj}_u \overline{AB} = A'B'.$$

Constructing the projection of the vector \overline{AB} on the axis u is illustrated in Fig. 83, where, for greater clarity, planes α and β have been drawn through the points A and B perpendicular to the axis u . The intersections of these planes with the axis u give the points A' and B' (since the planes α and β are perpendicular to the axis u , the straight lines AA' and BB' are also perpendicular to the axis u).

132. Take an arbitrary point S in space; from this point, draw a ray in the direction of the vector \overline{AB} , and another ray in the direction of the axis u (Fig. 83). The angle φ made by these two rays is called the *angle of inclination* of the vector \overline{AB} with respect to the axis u . It is obviously immaterial where to choose the point S for constructing the angle φ . It is equally obvious that the angle φ will remain unchanged if the axis u is replaced by another similarly directed axis. Let v denote the axis having the

same direction as u and passing through the point A . According to what has just been said, the angle of inclination of the vector \overline{AB} with respect to the axis v is equal to φ . Let C be the point where the axis v meets the plane β . Since the axis v is parallel to the axis u , which is perpendicular to the plane β , it follows that the axis v is also perpendicular to the plane β . Consequently, AC is the projection of the vector \overline{AB} on the axis v . Furthermore,

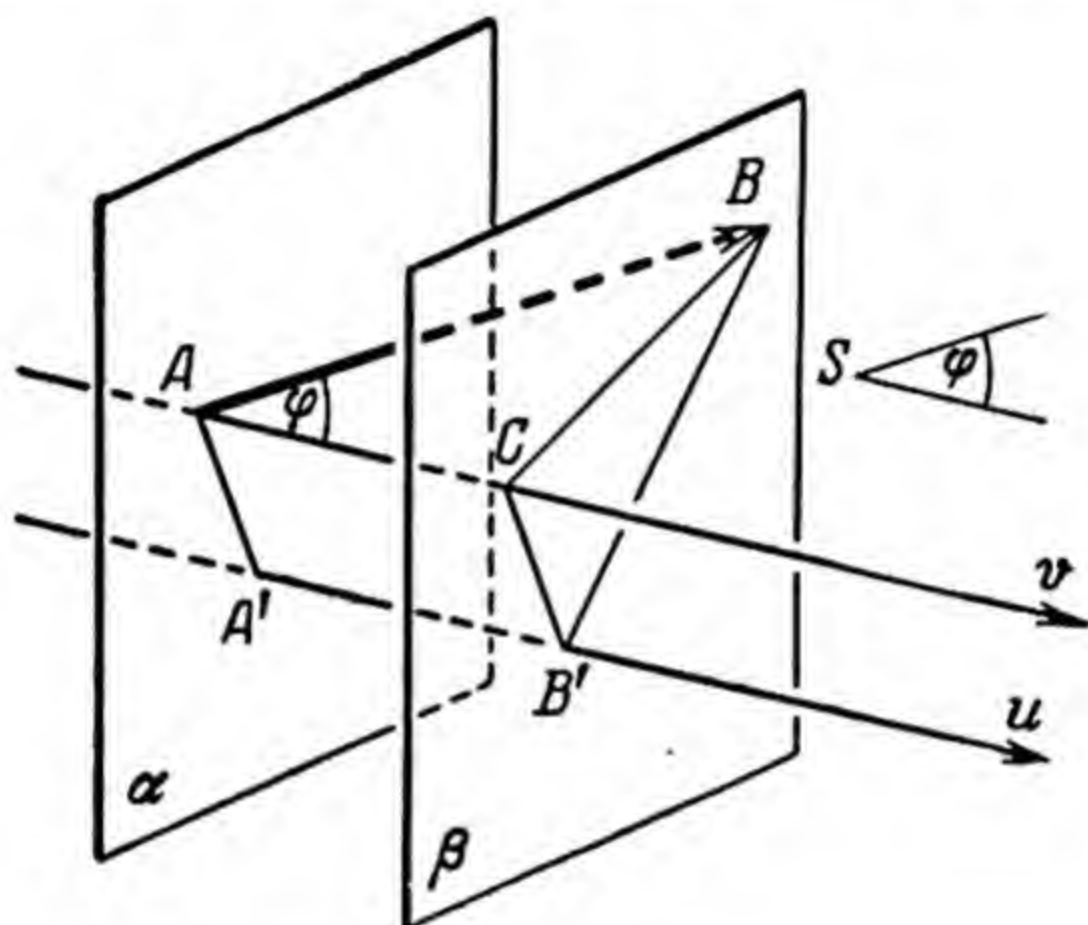


Fig. 83.

since the axes u and v are parallel and similarly directed, their segments included between the parallel planes α and β have equal values: $A'B' = AC$. Hence

$$\text{proj}_u \overline{AB} = \text{proj}_v \overline{AB}. \quad (1)$$

On the other hand, since the vector \overline{AB} and the axis v lie in the same plane, we may apply to them formula (7) of Art. 20, which gives

$$\text{proj}_v \overline{AB} = |\overline{AB}| \cos \varphi. \quad (2)$$

From (1) and (2), we obtain

$$\text{proj}_u \overline{AB} = |\overline{AB}| \cos \varphi. \quad (3)$$

If the vector \overline{AB} is denoted, for brevity, by the single letter a , formula (3) takes the form

$$\text{proj}_u a = |a| \cos \varphi. \quad (4)$$

Thus, the projection of a vector on an axis is equal to the product of the modulus of the vector and the cosine of its angle of inclination with respect to that axis.

133. Consider two equal vectors $\overline{A_1B_1}$, $\overline{A_2B_2}$ and an axis u . Since equal vectors have the same modulus and the same angle of inclination with respect to the axis u , it follows that the application of formula (3) to each of the vectors will yield the same result:

$$\text{proj}_u \overline{A_1B_1} = \text{proj}_u \overline{A_2B_2}.$$

Thus, equal vectors have their projections on the same axis equal.

§ 45. The Projections of a Vector on the Coordinate Axes

134. Assuming that a rectangular cartesian system of coordinates has been attached to space, let us consider an arbitrary vector \mathbf{a} . Let X denote the projection of the vector \mathbf{a} on the axis Ox , Y the projection of \mathbf{a} on the axis Oy , and Z the projection of \mathbf{a} on the axis Oz .

By Art. 133, every vector equal to \mathbf{a} has the same numbers X , Y , Z as its projections on the coordinate axes.

Conversely, if the projections of a vector \mathbf{b} on the coordinate axes are the same as those of a vector \mathbf{a} , then $\mathbf{b} = \mathbf{a}$. To verify this, draw both vectors \mathbf{a} and \mathbf{b} from the origin of coordinates and denote their terminal points by the letters A and B , respectively. Since the vectors \mathbf{a} and \mathbf{b} have the same projection X on the axis Ox , it is evident that the points A and B must lie in the same plane perpendicular to the axis Ox , namely, in the plane whose intercept on the axis Ox is equal in value to X . For a similar reason, the points A and B must lie in the same plane perpendicular to the axis Oy , namely, in the plane whose intercept on the axis Oy is equal to Y ; A and B must also lie in the same plane perpendicular to the axis Oz , namely, in the plane whose intercept on the axis Oz is equal to Z . But then the points A and B necessarily coincide, because the three above-mentioned planes intersect in a single point. Accordingly,

$$\mathbf{b} = \overline{OB} = \overline{OA} = \mathbf{a}.$$

This means that the projections of a vector on the coordinate axes completely determine it as a free vector, that is, to within its position in space. The projections X , Y , Z of a vector \mathbf{a} are therefore called its (cartesian) coordinates.

To express the fact that a vector \mathbf{a} has coordinates X, Y, Z , we shall henceforth use the notation

$$\mathbf{a} = \{X, Y, Z\},$$

where the right-hand member is considered as a new symbol for a vector.

135. In analytic geometry it is often necessary to calculate the coordinates of a vector, that is, its projections on the coordinate axes, from the given coordinates of the terminal and initial points of the vector. This problem is solved by the following

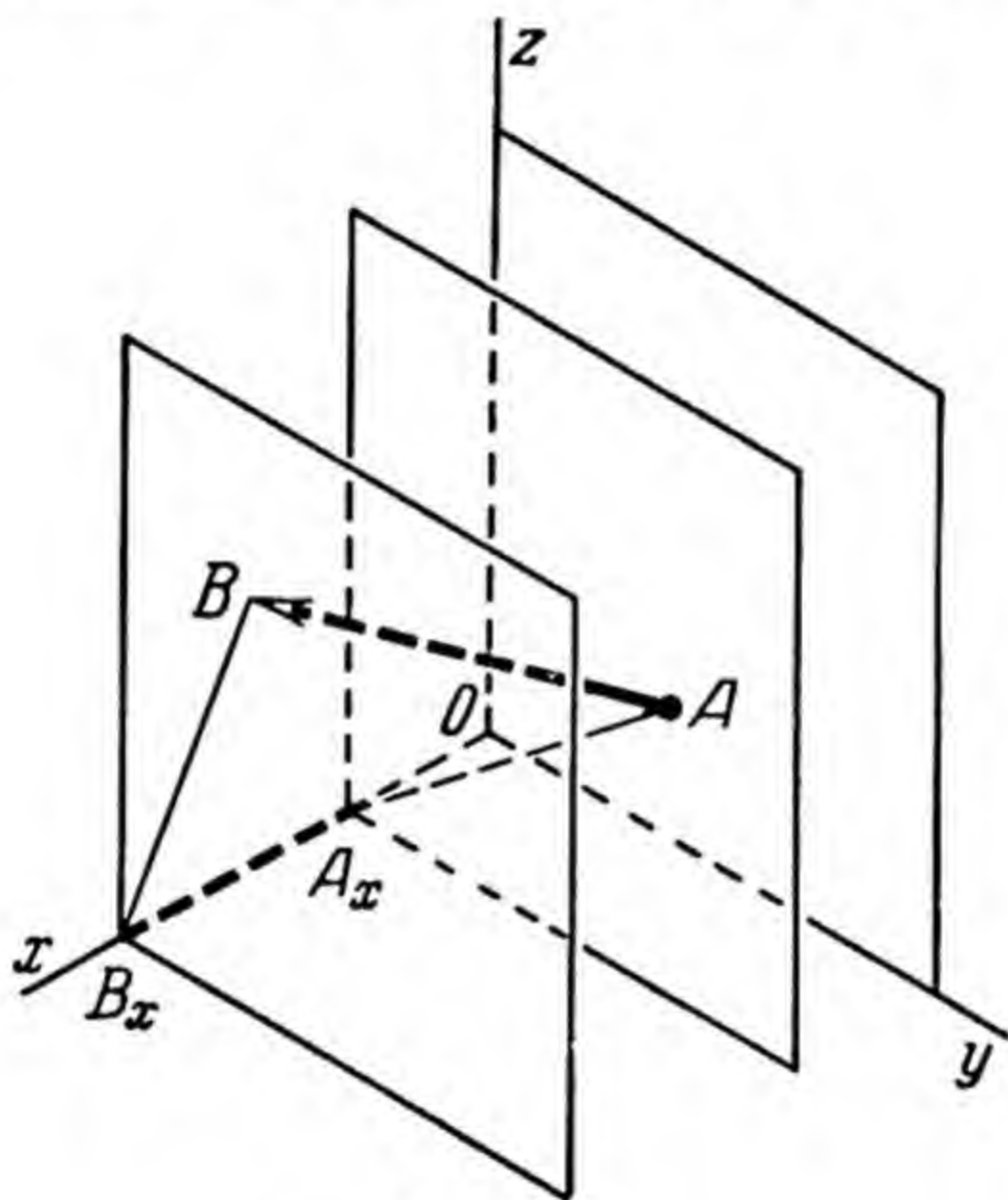


Fig. 84.

Theorem 15. For any two points $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$, the coordinates of the vector \overrightarrow{AB} are determined by the formulas

$$X = x_2 - x_1, \quad Y = y_2 - y_1, \quad Z = z_2 - z_1. \quad (1)$$

Proof. From the points A and B , drop perpendiculars on the axis Ox and denote their feet by A_x, B_x (see Fig. 84, where planes perpendicular to the axis Ox have been drawn through A and B , for greater clarity). The coordinates of the points A_x and B_x on the axis Ox are x_1 and x_2 , respectively. Hence, by Theorem 1 of Art. 5,

$$A_x B_x = x_2 - x_1.$$

But $A_x B_x = X$, so that $X = x_2 - x_1$. The relations $Y = y_2 - y_1$ and $Z = z_2 - z_1$ are established in a similar way.

Thus, to obtain the coordinates of a vector, subtract the coordinates of its initial point from the corresponding coordinates of its terminal point.

136. Let $M(x, y, z)$ be an arbitrary point in space. The vector $\mathbf{r} = \overline{OM}$ drawn from the origin of coordinates to the point M is called the radius vector of that point.

Calculating the coordinates of the vector \overline{OM} by formulas (1), that is, setting $x_2 = x$, $y_2 = y$, $z_2 = z$, $x_1 = 0$, $y_1 = 0$, $z_1 = 0$, we get

$$X = x, \quad Y = y, \quad Z = z,$$

which means that the coordinates of the radius vector of a point M are the same as the coordinates of that point. It should, however, be noted that, apart from formulas (1), this last statement follows immediately from the definition of the cartesian coordinates of a point M (see Art. 122).

137. Let there be given an arbitrary vector $\mathbf{a} = \{X, Y, Z\}$. We shall now derive a formula for computing the modulus of the vector \mathbf{a} from the known coordinates X, Y, Z of this vector.

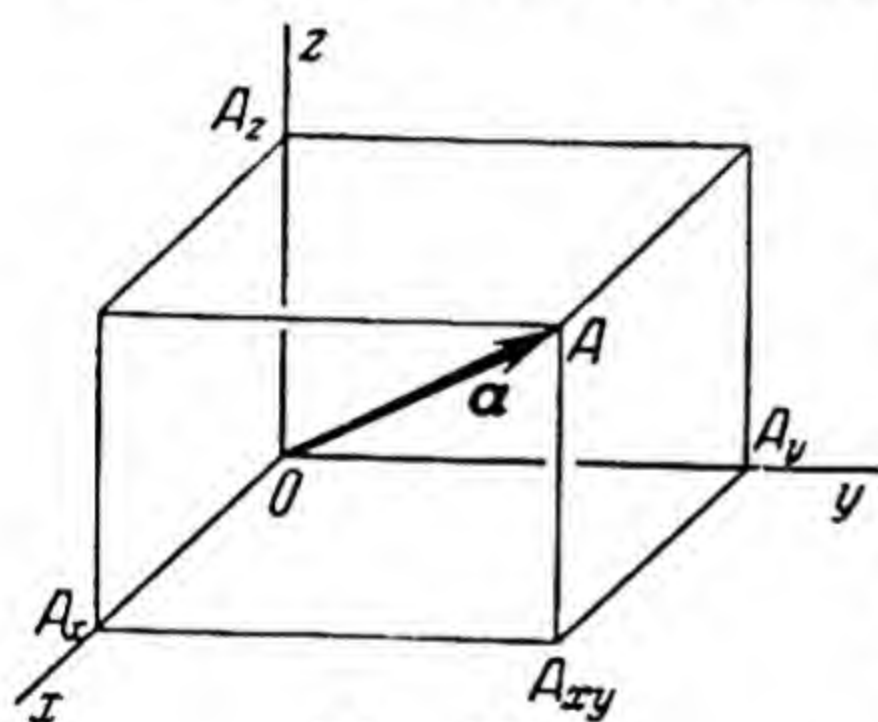


Fig. 85.

For simplicity, we shall assume that the vector is drawn from the origin of coordinates. Through the terminal end A of the vector \mathbf{a} , let three planes be passed perpendicular to the coordinate axes; denote the points where the planes intersect the axes by A_x, A_y, A_z , respectively. These planes, together with the coordinate planes, form a rectangular parallelepiped having the segment OA as a diagonal (Fig. 85). It is known from elementary geometry that the square of the length of the diagonal of a rectangular parallelepiped is equal to the sum of the squares of the lengths of its concurrent edges. Hence,

$$OA^2 = OA_x^2 + OA_y^2 + OA_z^2.$$

Now, $OA = |\mathbf{a}|$, $OA_x = X$, $OA_y = Y$, $OA_z = Z$; thus, the above relation becomes

$$|\mathbf{a}|^2 = X^2 + Y^2 + Z^2,$$

or

$$|a| = \sqrt{X^2 + Y^2 + Z^2}. \quad (2)$$

This is the desired expression for the modulus of an arbitrary vector in terms of its coordinates.

§ 46. Direction Cosines

138. Let α, β, γ denote the angles which a vector a makes with the coordinate axes; $\cos \alpha, \cos \beta, \cos \gamma$ are called *the direction cosines of the vector a* . They are called so because, once these cosines are given, they determine the direction of the vector.

If both the modulus and the direction cosines of a vector are given, the vector is thereby completely determined (as a free vector). In this case *the coordinates of the vector can be computed*, according to Art. 132, from the formulas

$$X = |a| \cos \alpha, \quad Y = |a| \cos \beta, \quad Z = |a| \cos \gamma. \quad (1)$$

139. We summarise the results of the two preceding articles in the form of the following

Theorem 16. For any vector a , its modulus $|a|$, direction cosines $\cos \alpha, \cos \beta, \cos \gamma$, and coordinates X, Y, Z are connected by the relations

$$X = |a| \cos \alpha, \quad Y = |a| \cos \beta, \quad Z = |a| \cos \gamma, \quad (1)$$

$$|a| = \sqrt{X^2 + Y^2 + Z^2}. \quad (2)$$

Note. The last four formulas enable us to calculate the direction cosines of a vector from the coordinates of this vector. For, from these formulas it follows that

$$\left. \begin{aligned} \cos \alpha &= \frac{X}{\sqrt{X^2 + Y^2 + Z^2}}, & \cos \beta &= \frac{Y}{\sqrt{X^2 + Y^2 + Z^2}}, \\ \cos \gamma &= \frac{Z}{\sqrt{X^2 + Y^2 + Z^2}}. \end{aligned} \right\} \quad (3)$$

Here the roots are taken in the arithmetical sense (as always in cases where no signs are prefixed to the radicals).

140. Squaring both members of each of relations (3) and adding, we find

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = \frac{X^2 + Y^2 + Z^2}{X^2 + Y^2 + Z^2};$$

hence

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1. \quad (4)$$

Relation (4) permits us to calculate any one of the angles α, β, γ , when the other two angles are given and it is also known whether the required angle is acute or obtuse.

§ 47. Distance Between Two Points. Division of a Line Segment in a Given Ratio

141. In solid analytic geometry (as well as in plane analytic geometry) every problem, however complex it may be, is reduced to certain elementary problems, such as the problem of determining the distance between two given points, the problem of dividing a line segment in a given ratio, the problem of calculating the angle between two vectors, etc. In the present section we shall take up the first two of these problems.

142. Given two arbitrary points $M_1 (x_1, y_1, z_1)$ and $M_2 (x_2, y_2, z_2)$, compute the distance d between them.

The desired result is obtained at once by the use of Theorem 15 (Art. 135) and formula (2) of the previous section.

In fact, we have

$$\overline{M_1 M_2} = \{x_2 - x_1, y_2 - y_1, z_2 - z_1\};$$

further, d is the modulus of the vector $\overline{M_1 M_2}$, so that

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \quad (1)$$

This formula furnishes the solution of the problem.

143. Given two arbitrary points $M_1 (x_1, y_1, z_1)$ and $M_2 (x_2, y_2, z_2)$, find on the line $M_1 M_2$ the point M dividing the segment $\overline{M_1 M_2}$ in the given ratio λ .

The solution of this problem is similar to that of the analogous problem in plane analytic geometry (see Art. 24). We shall therefore give the desired result directly: If x, y, z denote the coordinates of the required point M , then

$$x = \frac{x_1 + \lambda x_2}{1 + \lambda}, \quad y = \frac{y_1 + \lambda y_2}{1 + \lambda}, \quad z = \frac{z_1 + \lambda z_2}{1 + \lambda}.$$

In particular, the coordinates of the midpoint of a given segment are obtained from these relations by setting $\lambda = 1$:

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}, \quad z = \frac{z_1 + z_2}{2}.$$

144. In solving other elementary problems of solid analytic geometry, it is convenient to use certain special operations on vectors, known as addition of vectors, multiplication of vectors by numbers, scalar multiplication of vectors, and vector multiplication of vectors. The definitions and basic properties of these operations are presented in the next three chapters.

LINEAR OPERATIONS ON VECTORS

§ 48. Definitions of Linear Operations

145. The linear operations on vectors are those of vector addition and multiplication of vectors by numbers.

Definition of the sum of two vectors. Let \mathbf{a} and \mathbf{b} be two given vectors. Draw \mathbf{b} from the terminal point of \mathbf{a} ; then the sum $\mathbf{a} + \mathbf{b}$ is defined as the vector extending from the initial point of the vector \mathbf{a} to the terminal point of the vector \mathbf{b} .

The construction of the sum $\mathbf{a} + \mathbf{b}$ is shown in Fig. 86. The rule for addition of vectors contained in the above definition is customarily called the triangle rule.

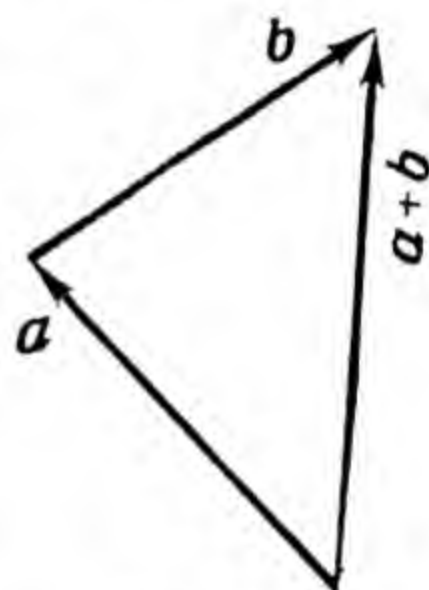


Fig. 86.

Note. When constructing the sum $\mathbf{a} + \mathbf{b}$ by the triangle rule, it may happen that the terminal point of \mathbf{b} will coincide with the initial point of \mathbf{a} ; then $\mathbf{a} + \mathbf{b}$ is a zero vector: $\mathbf{a} + \mathbf{b} = \mathbf{0}$.

Definition of the product of a vector and a number. Let there be given a vector \mathbf{a} and a number α , and let $|\mathbf{a}|$ and $|\alpha|$ denote their respective moduli. The product $\alpha\mathbf{a}$ (or $\mathbf{a}\alpha$) is defined as the vector collinear with \mathbf{a} , equal in length to $|\mathbf{a}| \cdot |\alpha|$ and having the same direction as the vector \mathbf{a} if $\alpha > 0$, or the opposite direction if $\alpha < 0$.

The operation of constructing the vector $\alpha\mathbf{a}$ is called the multiplication of the vector \mathbf{a} by the number α .

Note 1. If $\mathbf{a} = \mathbf{0}$ or $\alpha = 0$, the product has its modulus equal to zero and, hence, is a zero vector. In this case, the product $\alpha\mathbf{a}$ has no definite direction.

Note 2. The operation of multiplication of a vector by a number may be visualised as follows: Multiplying a vector \mathbf{a} by a number α means that the vector \mathbf{a} is "stretched α -fold". The use of the word "stretched" is, of course, a matter of pure convention; for instance, if $\alpha = \frac{1}{2}$, then "stretched α -fold" actually means

"shrunk by half"; if α is a negative number, "stretched α -fold" means that the vector is elongated $|\alpha|$ -fold (gets α times its former modulus), while its direction is reversed.

§ 49. Basic Properties of Linear Operations

146. We shall now establish the basic properties of the linear operations used in vector analysis.

First of all, we shall show that *the sum of any two vectors is independent of the order in which they are added.*

For this purpose, let us consider two arbitrary vectors \mathbf{a} and \mathbf{b} . Since geometric vectors are free vectors, we may draw \mathbf{a} and \mathbf{b} from a common initial point O chosen at will. Let the letters A and B denote the terminal points of the vectors \mathbf{a} and \mathbf{b} so drawn (Fig. 87). Next, apply the vector \mathbf{b} at the point A , denote its terminal point (in this new position) by the letter C , and join the points B and C by a segment. Clearly, the vector \overline{BC} has the same length and direction as the vector \overline{OA} ; hence $\overline{BC} = \mathbf{a}$.

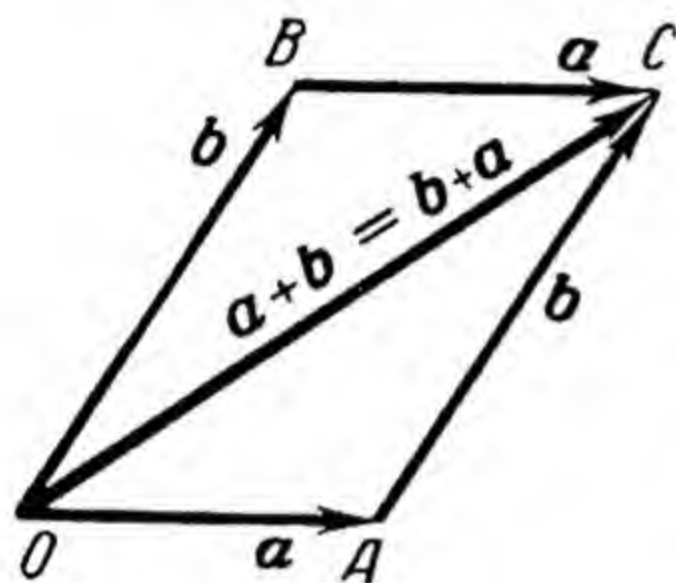


Fig. 87.

Considering now the geometric figure OAC and recalling the rule for addition of vectors (the triangle rule), we find that $\overline{OC} = \mathbf{a} + \mathbf{b}$. On the other hand, when considering the figure OBC , we find that, by the same rule, $\overline{OC} = \mathbf{b} + \mathbf{a}$. Hence

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (1)$$

as was to be shown.

The property expressed by identity (1) is called *the commutative property of vector addition.*

Note. The figure $OABC$ is referred to as *the parallelogram constructed on the vectors \mathbf{a} , \mathbf{b} with common initial point O* , and the vector \overline{OC} is said to be its diagonal (even if $\mathbf{a} = \overline{OA}$ and $\mathbf{b} = \overline{OB}$ lie on the same straight line, that is, if $OABC$ is not, properly speaking, a parallelogram). Accordingly, the rule for addition of vectors can now be given the following new wording:

If vectors \mathbf{a} and \mathbf{b} are drawn from a common initial point and a parallelogram is constructed on them, then the sum $\mathbf{a} + \mathbf{b}$ (or $\mathbf{b} + \mathbf{a}$) is that diagonal of the parallelogram which extends from the common initial point of \mathbf{a} and \mathbf{b} .

Expressed in this form, the rule for vector addition is called the parallelogram rule.

147. Now that we have defined the sum of two vectors, the definition of the sum of any number of vectors will follow naturally enough.

Let, for example, \mathbf{a} , \mathbf{b} , and \mathbf{c} be three given vectors. Adding \mathbf{a} and \mathbf{b} gives the vector $\mathbf{a} + \mathbf{b}$. Next, by adding \mathbf{c} to $\mathbf{a} + \mathbf{b}$ we obtain the vector $(\mathbf{a} + \mathbf{b}) + \mathbf{c}$. On the other hand, we can also construct the vector $\mathbf{a} + (\mathbf{b} + \mathbf{c})$ that is, add the sum $\mathbf{b} + \mathbf{c}$ to the vector \mathbf{a} .

It is easy to show that, for any three vectors \mathbf{a} , \mathbf{b} , \mathbf{c} , we always have the relation

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}). \quad (2)$$

The property expressed by identity (2) is called *the associative property* of vector addition.

To prove the associative property, draw the vector \mathbf{b} from the terminal point of the vector \mathbf{a} , and then draw \mathbf{c} from the terminal point of \mathbf{b} . With the vectors so drawn, denote the initial point of the vector \mathbf{a} by the letter O , the terminal point of \mathbf{a} by A , the terminal point of \mathbf{b} by B , and the terminal point of \mathbf{c} by C (Fig. 88). Then

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = (\overline{OA} + \overline{AB}) + \overline{BC} = \overline{OB} + \overline{BC} = \overline{OC},$$

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = \overline{OA} + (\overline{AB} + \overline{BC}) = \overline{OA} + \overline{AC} = \overline{OC}.$$

Consequently,

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}),$$

as was to be proved.

Owing to the associativity of vector addition, we may speak of the sum of three vectors \mathbf{a} , \mathbf{b} , \mathbf{c} and write it in the form $\mathbf{a} + \mathbf{b} + \mathbf{c}$ without having to specify whether $\mathbf{a} + \mathbf{b} + \mathbf{c} = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$ or $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ is meant. The sum of four, five or any other number of vectors may be defined in a similar way.

In practical addition of vectors, we can do without the consecutive construction of the intermediate sums; the sum of any number of vectors can be constructed at once by using the following

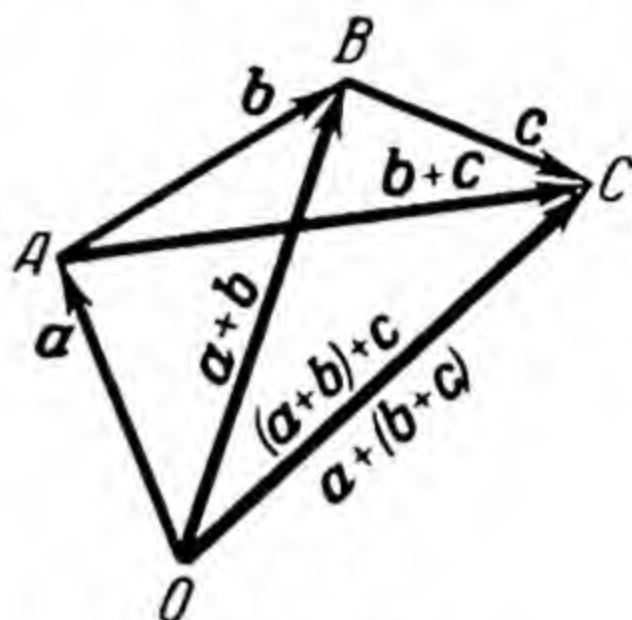


Fig. 88.

General rule for addition of vectors. To construct the sum of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$, draw \mathbf{a}_2 from the terminal point of \mathbf{a}_1 , next draw \mathbf{a}_3 from the terminal point of \mathbf{a}_2 , after this draw \mathbf{a}_4 from the terminal point of \mathbf{a}_3 , and so on, until the last vector \mathbf{a}_n is drawn. Then the sum $\mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_n$ will be the vector directed from the initial point of \mathbf{a}_1 to the terminal point of \mathbf{a}_n .

Let O be the initial point of the vector \mathbf{a}_1 , and let A_1, A_2, \dots, A_n denote the respective terminal ends of the vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ drawn in accordance with the rule just formulated. The figure $OA_1A_2 \dots A_n$ is referred to as the broken line with vector segments $\overline{OA_1} = \mathbf{a}_1, \overline{A_1A_2} = \mathbf{a}_2, \dots, \overline{A_{n-1}A_n} = \mathbf{a}_n$; the vector $\overline{OA_n}$ is said to close the polygon $OA_1A_2 \dots A_n$. Since

$$\begin{aligned}\overline{OA_n} &= \overline{OA_1} + \overline{A_1A_2} + \overline{A_2A_3} + \dots + \overline{A_{n-1}A_n} = \\ &= \mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_n,\end{aligned}$$

we may say that *the sum of several vectors is constructed by closing the polygon.* (Fig. 89 illustrates the construction of the sum of four vectors.)

Note. In Art. 146 it was established that the sum of two vectors is independent of the order in which they are added. From

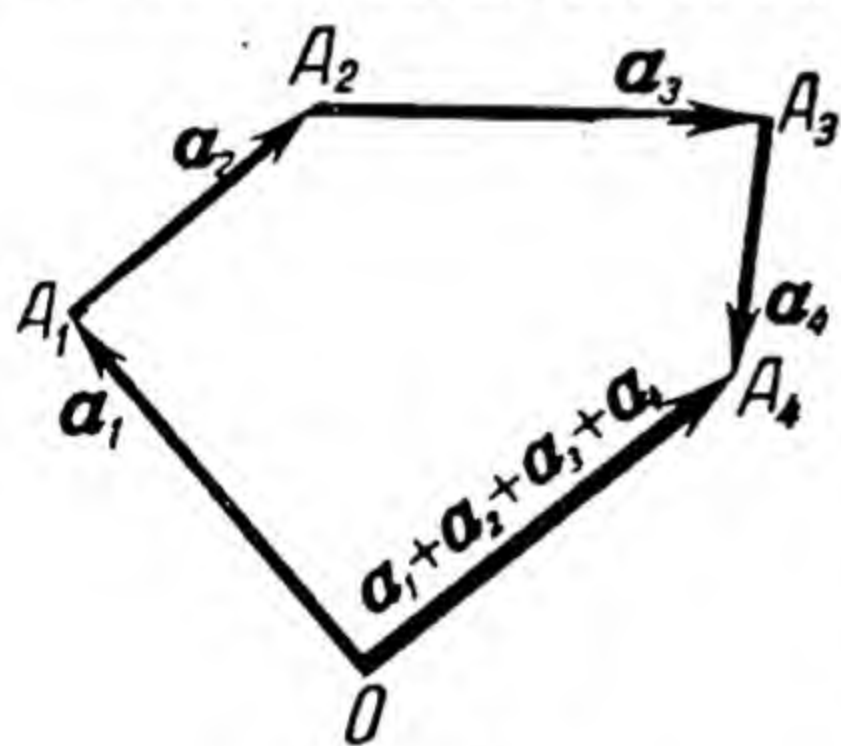


Fig. 89.

this and from the associativity of vector addition, it follows that *the sum of any number of vectors is also independent of the order in which the vectors are added.*

148. We shall now point out three properties of the linear operations, which refer both to addition of vectors and to multiplication of vectors by numbers. These properties are expressed by means of the following three

identities (in which λ and μ denote arbitrary numbers; \mathbf{a} and \mathbf{b} arbitrary vectors):

- (1) $(\lambda + \mu) \mathbf{a} = \lambda \mathbf{a} + \mu \mathbf{a},$
- (2) $\lambda (\mu \mathbf{a}) = (\lambda \mu) \mathbf{a},$
- (3) $\lambda (\mathbf{a} + \mathbf{b}) = \lambda \mathbf{a} + \lambda \mathbf{b}.$

The truth of the first identity becomes apparent when the identity is expressed in geometric terms: it means that stretch-

ing the vector \mathbf{a} $\lambda + \mu$ -fold gives the same vector as adding the vector \mathbf{a} stretched μ -fold to the vector \mathbf{a} stretched λ -fold *).

The geometric meaning of the second identity is equally transparent: stretching the vector \mathbf{a} μ -fold followed by stretching the vector $\mu\mathbf{a}$ λ -fold results in the same vector as stretching the vector \mathbf{a} $\lambda\mu$ -fold.

The third identity follows from the theory of similar figures. For, the vector $\mathbf{a} + \mathbf{b}$ forms a diagonal of the parallelogram constructed on the vectors \mathbf{a} and \mathbf{b} (assuming that \mathbf{a} and \mathbf{b} have been drawn from a common initial point); when the vectors \mathbf{a} , \mathbf{b} , and $\mathbf{a} + \mathbf{b}$ are all stretched λ -fold, the resulting figure is similar to the original parallelogram and is, therefore, another parallelogram. Accordingly, $\lambda(\mathbf{a} + \mathbf{b})$ is a diagonal of the parallelogram constructed on the vectors $\lambda\mathbf{a}$ and $\lambda\mathbf{b}$; hence $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$. (See Fig. 90, which corresponds to the case $\lambda > 0$; all vectors shown in the figure are assumed to issue from the point O .)

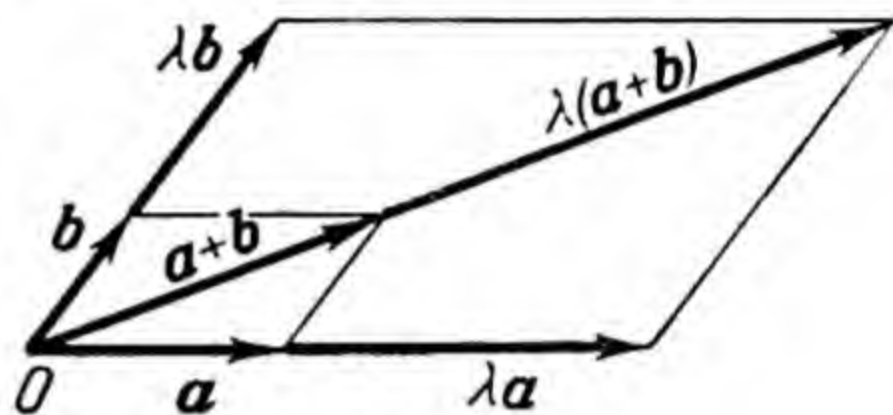


Fig. 90.

These three properties of the linear operations are of fundamental importance, as they enable us to carry out calculations in vector algebra in, basically, the same way as in ordinary algebra. The first of these properties expresses the permissibility of "distributing" the vector factor among the terms of the scalar factor; the third property expresses the permissibility of "distributing" the scalar factor among the components of the vector factor. Both properties are therefore called *distributive*. Together they permit us, when multiplying a scalar polynomial by a vector polynomial, to perform the operation "term by term".

The second property enables us to "associate" scalar factors into groups when multiplying a vector by several scalars in succession (for example, $2(5\mathbf{a}) = 10\mathbf{a}$). The second property is therefore called *associative*.

§ 50. The Vector Difference

149. Vector algebra includes subtraction of vectors; as in arithmetic, this operation is the inverse of addition.

Consider two arbitrary vectors \mathbf{a} and \mathbf{b} . The difference $\mathbf{b} - \mathbf{a}$

*) Here the term "stretching" is to be understood as specified in Note 2 of Art. 145.

is defined as the vector which, added to the vector \mathbf{a} , gives the vector \mathbf{b} .

Draw the vectors \mathbf{a} and \mathbf{b} from a common initial point O , and mark their terminal points as A and B (Fig. 91). Now let us find the difference $\mathbf{b} - \mathbf{a}$. Suppose that the desired vector $\mathbf{b} - \mathbf{a}$ is applied at the point A ; then its terminal point must coincide with the point B , since adding $\mathbf{b} - \mathbf{a}$ to the vector $\mathbf{a} = \overline{OA}$ must give the vector $\mathbf{b} = \overline{OB}$.

Hence, the difference $\mathbf{b} - \mathbf{a}$ is precisely the vector \overline{AB} :

$$\mathbf{b} - \mathbf{a} = \overline{AB}.$$

Thus, the difference of two vectors drawn from the same initial point is the vector extending from the terminal point of

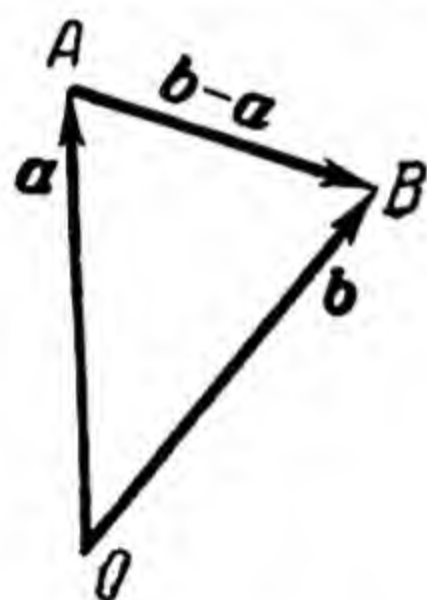


Fig. 91.

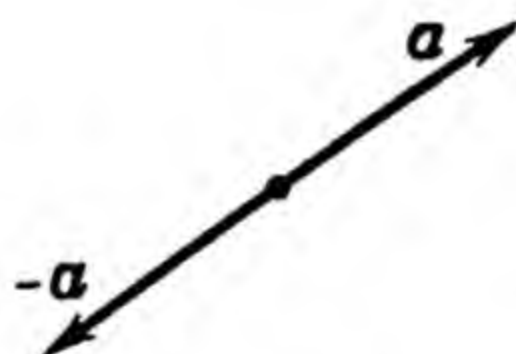


Fig. 92.

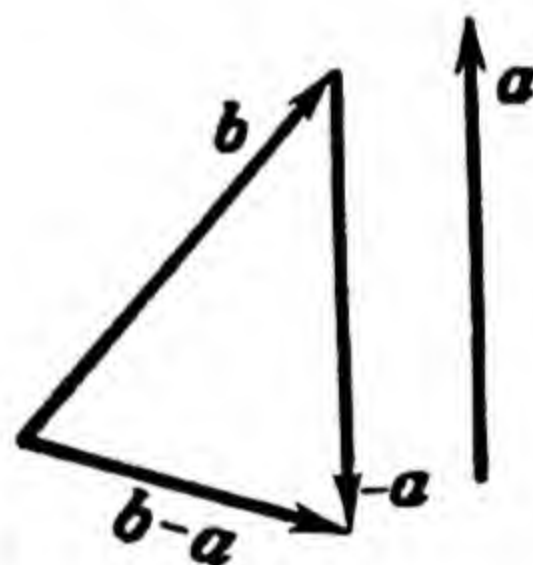


Fig. 93.

the vector "subtrahend" to the terminal point of the vector "minuend".

150. Besides the arbitrary vector \mathbf{a} , let us now consider the vector $(-1)\mathbf{a}$. The vector $(-1)\mathbf{a}$ is called the negative of the vector \mathbf{a} and is denoted by the symbol $-\mathbf{a}$:

$$-\mathbf{a} = (-1)\mathbf{a}.$$

Since multiplying the vector \mathbf{a} by -1 results in a vector which has the same modulus as, and is collinear with, the vector \mathbf{a} , but points in the opposite direction (Fig. 92), the vectors \mathbf{a} and $-\mathbf{a}$ are sometimes spoken of as *equal but opposite vectors*.

If the vector $-\mathbf{a}$ is drawn from the terminal point of the vector \mathbf{a} , then the terminal point of $-\mathbf{a}$ will coincide with the initial point of \mathbf{a} ; consequently, $\mathbf{a} + (-\mathbf{a})$ is a zero vector:

$$\mathbf{a} + (-\mathbf{a}) = \mathbf{0}. \quad (1)$$

151. We now return to the consideration of the two arbitrary vectors \mathbf{a} and \mathbf{b} . From identity (1), it immediately follows that

$$\mathbf{b} - \mathbf{a} = \mathbf{b} + (-\mathbf{a}). \quad (2)$$

For,

$$\mathbf{a} + [\mathbf{b} + (-\mathbf{a})] = \mathbf{b} + [\mathbf{a} + (-\mathbf{a})] = \mathbf{b} + 0 = \mathbf{b};$$

hence, the vector $\mathbf{b} + (-\mathbf{a})$, added to the vector \mathbf{a} , gives the vector \mathbf{b} , which means that the vector $\mathbf{b} + (-\mathbf{a})$ is the difference $\mathbf{b} - \mathbf{a}$.

Relation (2) expresses the following new rule for subtraction: *To obtain the difference $\mathbf{b} - \mathbf{a}$, add the negative of the vector \mathbf{a} to the vector \mathbf{b}* (see Fig. 93; from the diagram, it is immediately clear that the sum of the vectors \mathbf{a} and $\mathbf{b} + (-\mathbf{a})$ is the vector \mathbf{b}). This rule is particularly convenient to use when constructing the result of the addition and subtraction of several vectors; for example, to find $x = \mathbf{a} - \mathbf{b} - \mathbf{c} + \mathbf{d} - \mathbf{e}$, we have merely to construct the sum of the vectors \mathbf{a} , $-\mathbf{b}$, $-\mathbf{c}$, \mathbf{d} , $-\mathbf{e}$, as shown in Art. 147.

§ 51. Fundamental Theorems on Projections

152. In this article we shall establish two important theorems on the projections of vectors.

Theorem 17. *The projection of the sum of vectors on an axis is equal to the sum of their projections on this axis:*

$$\text{proj}_u(\mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_n) = \text{proj}_u \mathbf{a}_1 + \text{proj}_u \mathbf{a}_2 + \dots + \text{proj}_u \mathbf{a}_n.$$

Proof. Form the broken line with vector segments $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ (see Art. 147); that is, draw the vector \mathbf{a}_2 from the terminal point of the vector \mathbf{a}_1 , then draw \mathbf{a}_3 from the terminal point of \mathbf{a}_2 , etc., and finally draw the vector \mathbf{a}_n from the terminal point of the vector \mathbf{a}_{n-1} . With the vectors so drawn, denote the initial point of \mathbf{a}_1 by the letter O , the terminal point of \mathbf{a}_1 by A_1 , the terminal point of \mathbf{a}_2 by A_2 , etc. Then

$$\mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_n = \overline{OA_n}. \quad (1)$$

Projecting all these points O, A_1, A_2, \dots, A_n on an axis u and denoting their respective projections by $O', A'_1, A'_2, \dots, A'_n$ (see Fig. 94, which corresponds to $n = 3$), we obtain

$$O'A'_1 = \text{proj}_u \mathbf{a}_1, \quad A'_1A'_2 = \text{proj}_u \mathbf{a}_2, \quad \dots, \quad A'_{n-1}A'_n = \text{proj}_u \mathbf{a}_n. \quad (2)$$

On the other hand, in consequence of relation (1),

$$\text{proj}_u(\mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_n) = \text{proj}_u \overline{OA_n} = O'A'_n. \quad (3)$$

Now, for any position of the points $O', A'_1, A'_2, \dots, A'_n$ on the axis u , we have (by Art. 3) the identity

$$O'A'_n = O'A'_1 + A'_1A'_2 + A'_2A'_3 + \dots + A'_{n-1}A'_n. \quad (4)$$

From identity (4) and formulas (2) and (3), we find that

$$\text{proj}_u(\mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_n) = \text{proj}_u \mathbf{a}_1 + \text{proj}_u \mathbf{a}_2 + \dots + \text{proj}_u \mathbf{a}_n.$$

The theorem is proved.

Theorem 18. *When a vector is multiplied by a number, the projection of the vector is multiplied by the same number:*

$$\text{proj}_u \alpha \mathbf{a} = \alpha \text{proj}_u \mathbf{a}.$$

Proof. Draw the vector \mathbf{a} from any point O on the axis u , and denote the terminal point of \mathbf{a} by the letter A . Next, draw the

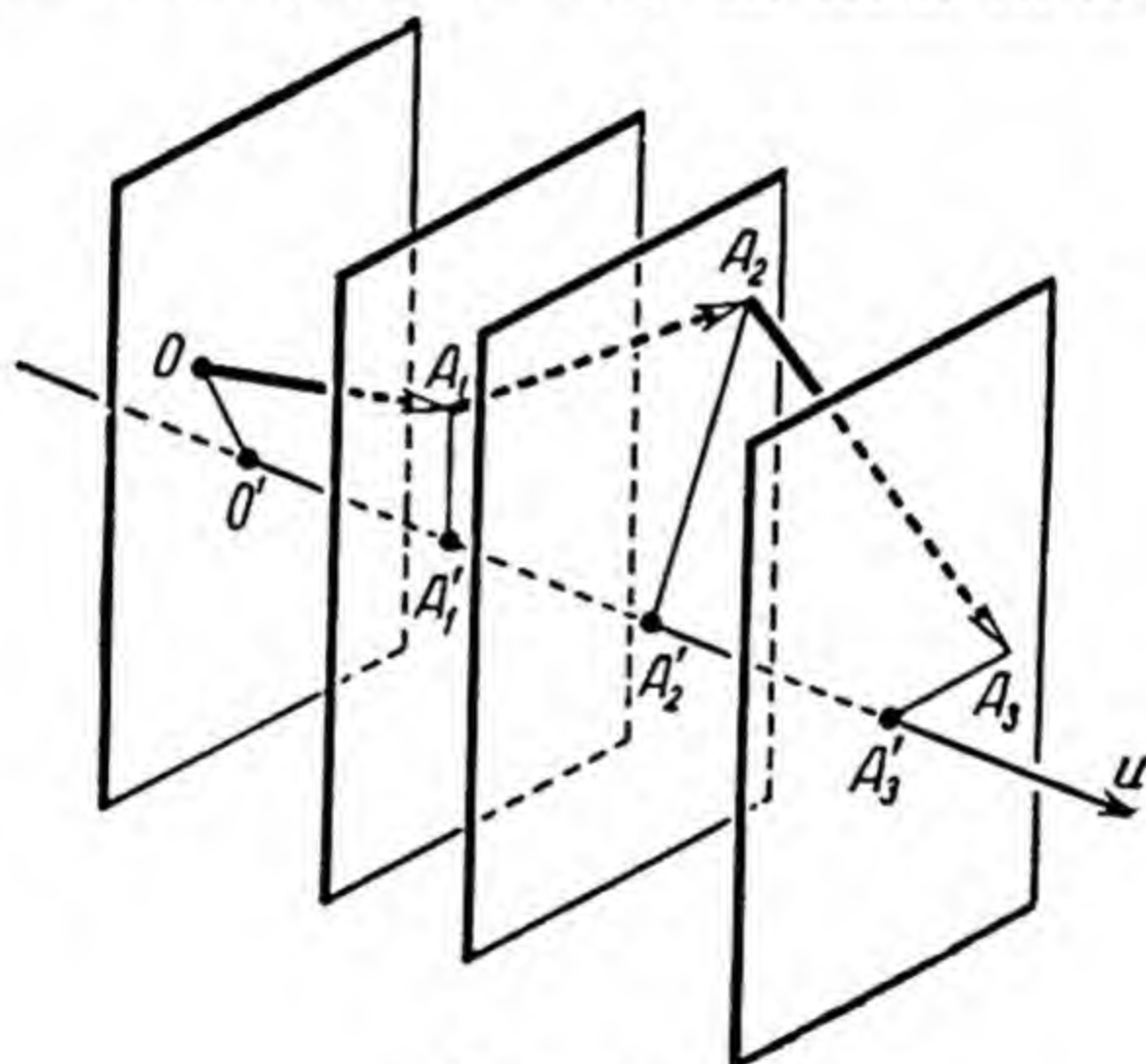


Fig. 94.

vector $\alpha \mathbf{a}$ from the same point O and mark the terminal point of $\alpha \mathbf{a}$ by B . Thus, $\overline{OA} = \mathbf{a}$, $\overline{OB} = \alpha \mathbf{a}$.

Consider the straight line v which contains the points O, A, B . Choose either of the directions of this line as the positive direction, thereby making the line an axis.

Project the points O, A and B upon the axis u ; let O', A', B' be their respective projections (Fig. 95a and b). By a well-known theorem of elementary geometry, we have the proportion

$$\left| \frac{O'B'}{O'A'} \right| = \left| \frac{OB}{OA} \right|. \quad (5)$$

If the segments \overline{OB} and \overline{OA} (which lie on the axis v) are like-directed, then the segments $\overline{O'B'}$ and $\overline{O'A'}$ (which lie on the axis u) are also like-directed; if, on the other hand, the segments \overline{OB} and \overline{OA} are oppositely directed, then the segments $\overline{O'B'}$ and $\overline{O'A'}$ also have opposite directions (this latter case, corresponding

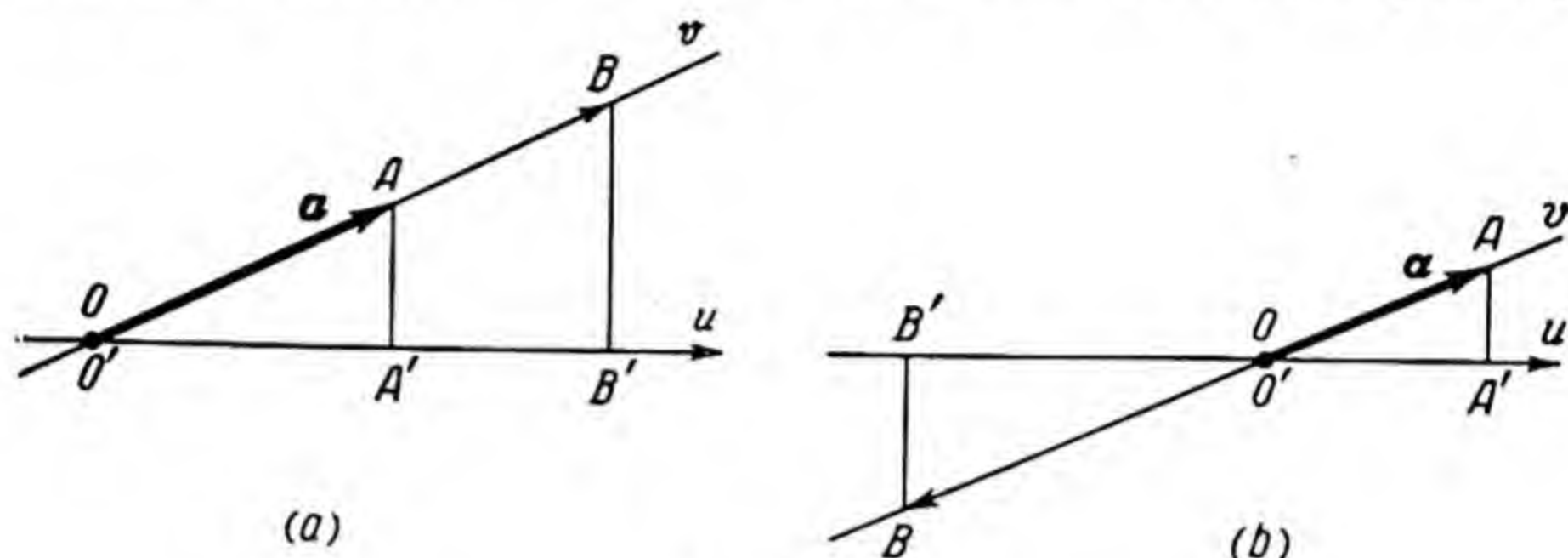


Fig. 95.

to a negative α , is illustrated in Fig. 95b). Thus, the ratios $\frac{O'B'}{O'A'}$ and $\frac{OB}{OA}$ of the *values* of these segments have *like signs*, so that relation (5) may be rewritten as

$$\frac{O'B'}{O'A'} = \frac{OB}{OA}.$$

Since $\overline{OB} = \alpha \mathbf{a}$ and $\overline{OA} = \mathbf{a}$, it follows that $\frac{OB}{OA} = \alpha$. Consequently, $\frac{O'B'}{O'A'} = \alpha$, or $O'B' = \alpha \cdot O'A'$. Hence

$$\text{proj}_u \alpha \mathbf{a} = \alpha \text{proj}_u \mathbf{a}.$$

The theorem is thus proved.

Note. This last theorem can be expressed in more graphic language as follows: *When a vector is stretched α -fold, its projection is also stretched α -fold.*

153. In Art. 134 we established the principle of determining every free vector in space by the specification of three numbers—the coordinates of the vector. It is essential for us to know what *arithmetical operations on the vector coordinates correspond to the linear operations performed on the vectors themselves*. This question is immediately resolved by Theorems 17 and 18 of Art. 152, once it is recalled that the (cartesian) coordinates of a vector are its projections on the coordinate axes. For, Theorem 17 means

that, when vectors are added, their coordinates are added. Thus, if

$$\mathbf{a} = \{X_1, Y_1, Z_1\} \quad \text{and} \quad \mathbf{b} = \{X_2, Y_2, Z_2\},$$

then

$$\mathbf{a} + \mathbf{b} = \{X_1 + X_2, Y_1 + Y_2, Z_1 + Z_2\}.$$

It follows that

$$\mathbf{a} - \mathbf{b} = \{X_1 - X_2, Y_1 - Y_2, Z_1 - Z_2\}.$$

These results may also be expressed symbolically by the single relation

$$\{X_1, Y_1, Z_1\} \pm \{X_2, Y_2, Z_2\} = \{X_1 \pm X_2, Y_1 \pm Y_2, Z_1 \pm Z_2\}. \quad (6)$$

Further, according to Theorem 18, when a vector is multiplied by a number, its coordinates are multiplied by the same number. Thus, if $\mathbf{a} = \{X, Y, Z\}$, then, for any number α ,

$$\alpha \mathbf{a} = \{\alpha X, \alpha Y, \alpha Z\}.$$

Another symbolic expression for this is

$$\alpha \{X, Y, Z\} = \{\alpha X, \alpha Y, \alpha Z\}. \quad (7)$$

154. From the foregoing, it is easy to deduce the condition for the collinearity of two vectors, whose coordinates are given.

For, vectors $\mathbf{a} = \{X_1, Y_1, Z_1\}$ and $\mathbf{b} = \{X_2, Y_2, Z_2\}$ are collinear if, and only if, one of them can be obtained by multiplying the other by some number: $\mathbf{b} = \lambda \mathbf{a}$ (we assume $\mathbf{a} \neq 0$). The vector relation

$$\mathbf{b} = \lambda \mathbf{a}$$

is equivalent to the three scalar relations

$$X_2 = \lambda X_1, \quad Y_2 = \lambda Y_1, \quad Z_2 = \lambda Z_1,$$

which mean that the coordinates of the vector \mathbf{b} are proportional to the coordinates of the vector \mathbf{a} . Consequently, the vectors $\mathbf{a} = \{X_1, Y_1, Z_1\}$ and $\mathbf{b} = \{X_2, Y_2, Z_2\}$ are collinear if, and only if,

$$\frac{X_2}{X_1} = \frac{Y_2}{Y_1} = \frac{Z_2}{Z_1},$$

that is, if their coordinates are proportional.

§ 52. Resolution of Vectors into Components

155. Assuming that a rectangular cartesian system of coordinates has been attached to space, we shall now consider, in connection with this system, a definite triad of vectors denoted by the symbols $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and determined by the following conditions:

(1) the vector i lies on the axis Ox , the vector j lies on the axis Oy , and the vector k lies on the axis Oz ;

(2) each of the vectors i, j, k points in the positive direction of the axis on which it lies;

(3) i, j, k are unit vectors, that is, $|i| = 1, |j| = 1, |k| = 1$.

We shall now show that *any vector in space may be expressed in terms of the vectors i, j, k by means of the linear operations.*

Consider an arbitrary vector a . We shall assume (for convenience in exposition) that the vector a is drawn from the origin of coordinates. Let A denote the terminal point of a . Draw a straight line through the point A parallel to the axis Oz . The line must intersect the plane Oxy ; denote the point of intersection by B . Next, through the point B draw a line parallel to the axis Oy , and a second line parallel to the axis Ox . The first of these lines must intersect the axis Ox , and the second line the axis Oy . Denote these points of intersection by A_x and A_y , respectively. Finally, draw a line through the point A parallel to the line OB ; the line so drawn must meet the axis Oz in some point, which will be denoted by A_z (Fig. 96).

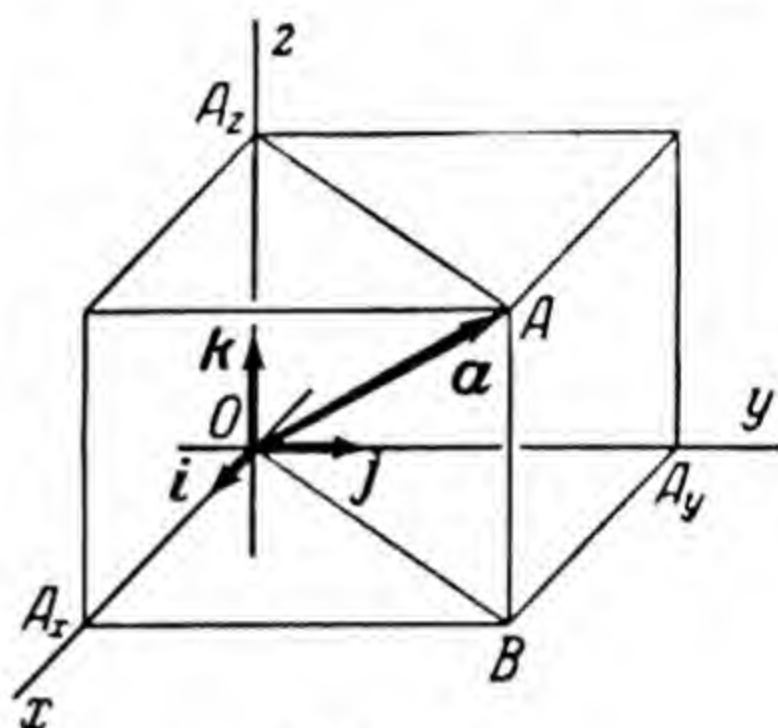


Fig. 96.

According to the rule for vector addition (applied to the parallelogram $OBAA_z$), we have

$$a = \overline{OB} + \overline{OA_z}. \quad (1)$$

Similarly, by applying the rule for vector addition to the parallelogram OA_yBA_x , we obtain

$$\overline{OB} = \overline{OA_x} + \overline{OA_y}. \quad (2)$$

From (1) and (2),

$$a = \overline{OA_x} + \overline{OA_y} + \overline{OA_z}. \quad (3)$$

Since the vectors $\overline{OA_x}$ and i lie on the same straight line, $\overline{OA_x}$ can be obtained by "stretching" the vector i ; hence we may write: $\overline{OA_x} = \lambda i$, where λ is some number.

In like manner, $\overline{OA_y} = \mu j$ and $\overline{OA_z} = \nu k$ (Fig. 96 corresponds to the case where the numbers λ, μ and ν are all positive).

From (3) and the relations $\overline{OA}_x = \lambda \mathbf{i}$, $\overline{OA}_y = \mu \mathbf{j}$, $\overline{OA}_z = \nu \mathbf{k}$, we obtain

$$\mathbf{a} = \lambda \mathbf{i} + \mu \mathbf{j} + \nu \mathbf{k}. \quad (4)$$

We have thus shown that any vector \mathbf{a} in space can actually be expressed in terms of the vectors \mathbf{i} , \mathbf{j} , \mathbf{k} by using the linear operations.

We shall refer to the triad of vectors \mathbf{i} , \mathbf{j} , \mathbf{k} as the *coordinate basis*, our purpose being to express all vectors in space in terms of this basis (as shown above).

The representation of a vector \mathbf{a} as the sum $\lambda \mathbf{i} + \mu \mathbf{j} + \nu \mathbf{k}$ is called the *resolution of a vector \mathbf{a} into components with respect to the basis \mathbf{i} , \mathbf{j} , \mathbf{k}* . The numbers λ , μ , ν are called the *coefficients* of this resolution; the vectors $\lambda \mathbf{i}$, $\mu \mathbf{j}$, $\nu \mathbf{k}$ are said to be the *components* of the vector \mathbf{a} with respect to the basis \mathbf{i} , \mathbf{j} , \mathbf{k} .

The vectors $\lambda \mathbf{i}$, $\mu \mathbf{j}$, $\nu \mathbf{k}$ are termed the *components* of \mathbf{a} because, by addition, they are compounded into the vector \mathbf{a} .

156. We shall now find out the geometric meaning of the coefficients λ , μ , ν . Since $\overline{OA}_x = \lambda \mathbf{i}$ and since \mathbf{i} is a unit vector, it follows that the number λ is the ratio of the segment \overline{OA}_x to the unit of measure taken with a plus or minus sign according as the segment agrees or disagrees in direction with the vector \mathbf{i} . In other words, λ is the *value* (understood as defined in Art. 2) of the segment \overline{OA}_x on the axis Ox , that is $\lambda = OA_x$. But OA_x is nothing more than the projection of the vector $\mathbf{a} = \overline{OA}$ on the axis Ox . Consequently,

$$\lambda = \text{proj}_x \mathbf{a} = X.$$

Similarly,

$$\mu = \text{proj}_y \mathbf{a} = Y, \quad \nu = \text{proj}_z \mathbf{a} = Z.$$

157. All that has been stated in Arts 155 and 156 may be summarised as the following

Theorem 19. Any vector \mathbf{a} can always be resolved into components with respect to the basis \mathbf{i} , \mathbf{j} , \mathbf{k} , that is, can always be expressed in the form

$$\mathbf{a} = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k};$$

the coefficients of this resolution are uniquely determined by the vector \mathbf{a} ; namely, X , Y , Z are the projections of the vector \mathbf{a} on the coordinate axes (that is, the coordinates of the vector \mathbf{a}).

158. Now, vectors can also be resolved into components with respect to bases other than the basis \mathbf{i} , \mathbf{j} , \mathbf{k} .

Let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ be three given vectors. For greater clarity, we shall assume that they are drawn from a common initial point O . No special conditions will be imposed upon these vectors (which may therefore have any lengths and make any angles with one another), the only requirement being that the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, when drawn from a common initial point O , must not lie in the same plane. With this proviso, the following theorem holds true:

Any vector \mathbf{a} can always be expressed as a linear combination of the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$:

$$\mathbf{a} = \lambda \mathbf{a}_1 + \mu \mathbf{a}_2 + \nu \mathbf{a}_3. \quad (5)$$

Such an expression for a vector \mathbf{a} is called its resolution with respect to the basis $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

To prove this theorem, let three axes Ox, Oy, Oz be drawn through the point O in the directions of the vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, respectively. Relation (5) can then be established by repeating verbally the reasoning and procedure used to derive (4), except that the vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ must everywhere be replaced by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ (and that the parallelepiped shown in Fig. 96 should be visualised as an oblique one).

It remains to find out the geometric meaning of the coefficients λ, μ, ν in (5). We have $\overline{OA_x} = \lambda \mathbf{a}_1$; hence λ is the value of the segment $\overline{OA_x}$ of the axis Ox , provided that the vector \mathbf{a}_1 has been chosen as the unit segment on this axis. Similar are the interpretations of μ and ν . The segments OA_x, OA_y, OA_z are sometimes called the oblique projections of the vector \mathbf{a} on the axes Ox, Oy, Oz . Accordingly, we may say that the coefficients λ, μ, ν in (5) are the values of the oblique projections of \mathbf{a} on the axes Ox, Oy, Oz , provided that each of these projections is measured on its axis in the appropriate scale. It follows that *the coefficients of the resolution of a given vector with respect to a given basis are determined uniquely* (since they represent completely specified geometric quantities).

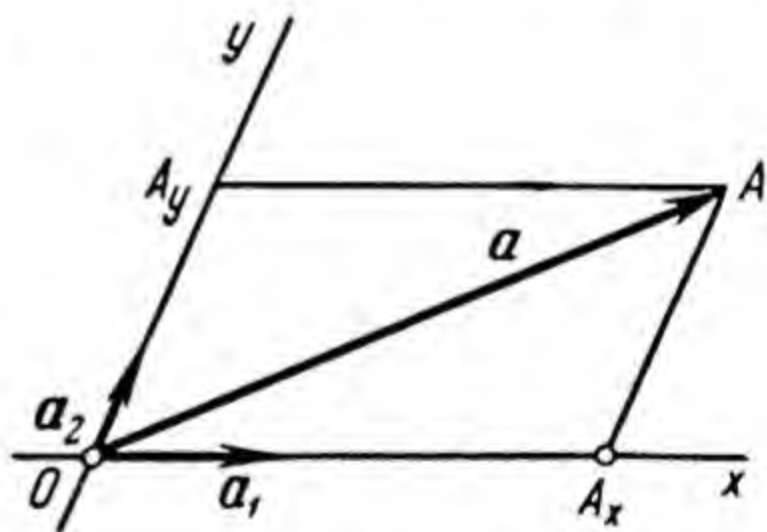


Fig. 96a.

159. If a vector \mathbf{a} lies in the plane of the vectors $\mathbf{a}_1, \mathbf{a}_2$, then its resolution with respect to the basis $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ has the form

$$\mathbf{a} = \lambda \mathbf{a}_1 + \mu \mathbf{a}_2,$$

that is, $v = 0$. For, the point A coincides in this case with the point B , and hence $\overline{OA}_z = 0$.

If we intend to consider vectors lying in one definite plane only, and to resolve them into components, it is sufficient to use for the purpose a basis formed by two vectors $\mathbf{a}_1, \mathbf{a}_2$, which lie in the given plane (the third vector becoming superfluous). Any pair of vectors $\mathbf{a}_1, \mathbf{a}_2$ in this plane may be taken as a basis, the only proviso being that the vectors $\mathbf{a}_1, \mathbf{a}_2$, when drawn from a common initial point O , must not be on the same straight line. In other words, the vectors forming a plane basis must not be collinear. Naturally, it is simpler to construct components in a plane than in space; the process is illustrated in Fig. 96a. We have

$$\mathbf{a} = \overline{OA} = \overline{OA}_x + \overline{OA}_y = \lambda \mathbf{a}_1 + \mu \mathbf{a}_2.$$

The segments OA_x, OA_y are the oblique projections of the vector \mathbf{a} upon the axes Ox, Oy ; the coefficients λ and μ are the values of these projections, provided that the vectors \mathbf{a}_1 and \mathbf{a}_2 have been adopted as the unit segments on the respective axes.

THE SCALAR PRODUCT OF VECTORS

§ 53. The Scalar Product and Its Basic Properties

160. *The scalar product of two vectors is defined as the number equal to the product of the moduli of these vectors by the cosine of their included angle.*

The scalar product of two vectors \mathbf{a} , \mathbf{b} is denoted by the symbol \mathbf{ab} .

Designating the angle between the vectors as φ , we may express the scalar product of vectors \mathbf{a} , \mathbf{b} by the formula

$$\mathbf{ab} = |\mathbf{a}| |\mathbf{b}| \cos \varphi. \quad (1)$$

For our subsequent work, it is important to note that $|\mathbf{b}| \cos \varphi = \text{proj}_{\mathbf{a}} \mathbf{b}$ (see Art. 132) and, consequently,

$$\mathbf{ab} = |\mathbf{a}| \text{proj}_{\mathbf{a}} \mathbf{b}. \quad (2)$$

Similarly, $|\mathbf{a}| \cos \varphi = \text{proj}_{\mathbf{b}} \mathbf{a}$, so that we also obtain

$$\mathbf{ab} = |\mathbf{b}| \text{proj}_{\mathbf{b}} \mathbf{a}. \quad (3)$$

Thus, the scalar product of vectors \mathbf{a} , \mathbf{b} may be regarded either as the product of two numbers, of which one is the modulus of \mathbf{a} and the other is the projection of \mathbf{b} upon the axis of \mathbf{a} , or as the product of two numbers, of which one is the modulus of \mathbf{b} and the other is the projection of \mathbf{a} upon the axis of \mathbf{b} .

161. The concept of the scalar product has its origin in mechanics; namely, if a (free) vector \mathbf{a} represents a force whose point of application experiences a displacement from the initial to the terminal point of a vector \mathbf{b} , then the work w done by this force is determined by the relation

$$w = |\mathbf{a}| |\mathbf{b}| \cos \varphi.$$

In vector analysis, this quantity is termed a *product* of two vectors because it possesses some algebraic properties of the ordinary product of numbers (these properties are dealt with in the next article); it is termed the *scalar* product because it is a scalar (i. e., a number).

162. The algebraic properties of the scalar product are as follows:

1. *The scalar product is commutative:*

$$ab = ba.$$

Proof. By definition, $ab = |a||b| \cos \varphi$ and $ba = |b||a| \cos \varphi$; but $|a||b| = |b||a|$ as an ordinary product of numbers, and hence $ab = ba$.

2. *The scalar product is associative with respect to multiplication by numbers:*

$$(\alpha a)b = \alpha(ab).$$

Proof. By formula (3), we have

$$(\alpha a)b = |b| \operatorname{proj}_b(\alpha a).$$

Now, according to Theorem 18 (Art. 152), $\operatorname{proj}_b(\alpha a) = \alpha \operatorname{proj}_b a$. Hence

$$(\alpha a)b = |b| \operatorname{proj}_b(\alpha a) = |b| \alpha \operatorname{proj}_b a = \alpha (|b| \operatorname{proj}_b a).$$

On the other hand, by the same formula (3), $|b| \operatorname{proj}_b a = ab$. Thus,

$$(\alpha a)b = \alpha (|b| \operatorname{proj}_b a) = \alpha(ab).$$

Note 1. From the properties 1 and 2, it follows that

$$(\alpha a)(\beta b) = (\alpha\beta)(ab).$$

For,

$$(\alpha a)(\beta b) = \alpha(a(\beta b)) = \alpha((\beta b)a) = \alpha(\beta(ba)) = (\alpha\beta)(ba) = (\alpha\beta)(ab).$$

3. *The scalar product is distributive with respect to addition:*

$$a(b+c) = ab + ac.$$

Proof. By formula (2), we have

$$a(b+c) = |a| \operatorname{proj}_a(b+c).$$

But, according to Theorem 17 (Art. 152), $\operatorname{proj}_a(b+c) = \operatorname{proj}_a b + \operatorname{proj}_a c$. Hence

$$\begin{aligned} a(b+c) &= |a| \operatorname{proj}_a(b+c) = |a|(\operatorname{proj}_a b + \operatorname{proj}_a c) = \\ &= |a| \operatorname{proj}_a b + |a| \operatorname{proj}_a c. \end{aligned}$$

On the other hand, by the same formula (2), $|a| \operatorname{proj}_a b = ab$ and $|a| \operatorname{proj}_a c = ac$. Consequently,

$$a(b+c) = |a| \operatorname{proj}_a b + |a| \operatorname{proj}_a c = ab + ac.$$

The last of the properties just established enables us to carry out the scalar multiplication of vector polynomials term by term.

In virtue of the first property, the order in which the factors are multiplied is immaterial. The second property of the scalar product permits us (see Note 1 above) to group together the scalar coefficients of the vector factors. For example,

$$\begin{aligned}(2\mathbf{a} + 5\mathbf{b})(3\mathbf{c} + 4\mathbf{d}) &= (2\mathbf{a} + 5\mathbf{b})(3\mathbf{c}) + (2\mathbf{a} + 5\mathbf{b})(4\mathbf{d}) = \\ &= (2\mathbf{a})(3\mathbf{c}) + (5\mathbf{b})(3\mathbf{c}) + (2\mathbf{a})(4\mathbf{d}) + (5\mathbf{b})(4\mathbf{d}) = \\ &= 6\mathbf{ac} + 15\mathbf{bc} + 8\mathbf{ad} + 20\mathbf{bd}.\end{aligned}$$

Note 2. In one respect, the scalar product of vectors *differs* essentially from the ordinary product of numbers; namely, since the scalar product of two vectors is a number, and no longer a vector, it is meaningless to speak about the scalar product of three or more vectors. Note that the symbol $(\mathbf{ab})\mathbf{c}$ can only be understood thus: $(\mathbf{ab})\mathbf{c}$ is the product of the number \mathbf{ab} and the vector \mathbf{c} , that is, $(\mathbf{ab})\mathbf{c}$ is the vector \mathbf{c} "stretched \mathbf{ab} -fold".

163. In this article we shall state a number of important geometric properties of the scalar product.

1. *If non-zero vectors \mathbf{a} and \mathbf{b} make an acute angle, the scalar product \mathbf{ab} is positive.*

For, if φ is an acute angle, then $\cos \varphi > 0$; consequently,

$$\mathbf{ab} = |\mathbf{a}| |\mathbf{b}| \cos \varphi > 0.$$

2. *If non-zero vectors \mathbf{a} and \mathbf{b} make an obtuse angle, the scalar product \mathbf{ab} is negative.*

For, if φ is an obtuse angle, then $\cos \varphi < 0$; consequently,

$$\mathbf{ab} = |\mathbf{a}| |\mathbf{b}| \cos \varphi < 0.$$

3. *If vectors \mathbf{a} and \mathbf{b} are mutually perpendicular, then $\mathbf{ab} = 0$.*

For, if \mathbf{a} and \mathbf{b} are perpendicular, then $\varphi = \frac{\pi}{2}$ and $\cos \varphi = 0$; hence

$$\mathbf{ab} = |\mathbf{a}| |\mathbf{b}| \cos \varphi = 0.$$

4. *If the scalar product of two vectors \mathbf{a} , \mathbf{b} is zero, then the vectors \mathbf{a} , \mathbf{b} are mutually perpendicular.*

For, if at least one of the vectors \mathbf{a} , \mathbf{b} is zero, it can be considered perpendicular to the other vector, since a zero vector may be regarded as having any direction; if, on the other hand, neither of the vectors is zero, then (in consequence of the relation $\mathbf{ab} = |\mathbf{a}||\mathbf{b}| \cos \varphi = 0$) $\cos \varphi = 0$, that is, $\varphi = \frac{\pi}{2}$, which means that the vectors \mathbf{a} and \mathbf{b} are perpendicular.

We may combine the last two properties into the following single statement: *Two vectors have their scalar product equal to zero if, and only if, they are mutually perpendicular.*

Finally, we shall point out one more property of the scalar product:

5. *The scalar product of a vector with itself is equal to the square of its modulus:*

$$\mathbf{a}\mathbf{a} = |\mathbf{a}|^2.$$

For, $\mathbf{a}\mathbf{a} = |\mathbf{a}||\mathbf{a}| \cos 0$; but $\cos 0 = 1$, and hence $\mathbf{a}\mathbf{a} = |\mathbf{a}|^2$.

Note. The scalar product $\mathbf{a}\mathbf{a}$ is called the scalar square of the vector \mathbf{a} and is denoted by the symbol \mathbf{a}^2 . From the foregoing, we have $\mathbf{a}^2 = |\mathbf{a}|^2$; that is, *the scalar square of a vector is equal to the square of its modulus.*

§ 54. Representation of the Scalar Product in Terms of the Coordinates of the Vector Factors

164. The following theorem makes it possible to compute the scalar product of two vectors from their coordinates, that is, from their projections on the axes of a rectangular cartesian coordinate system.

Theorem 20. *Given the coordinates of vectors \mathbf{a} and \mathbf{b} :*

$$\mathbf{a} = \{X_1, Y_1, Z_1\}, \quad \mathbf{b} = \{X_2, Y_2, Z_2\},$$

their scalar product is determined by the formula

$$\mathbf{a}\mathbf{b} = X_1X_2 + Y_1Y_2 + Z_1Z_2.$$

Proof. We shall first draw up a "multiplication table" for the base vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$:

$$\left. \begin{array}{lll} \mathbf{i}^2 = 1, & \mathbf{ij} = 0, & \mathbf{ik} = 0, \\ \mathbf{ji} = 0, & \mathbf{j}^2 = 1, & \mathbf{jk} = 0, \\ \mathbf{ki} = 0, & \mathbf{kj} = 0, & \mathbf{k}^2 = 1. \end{array} \right\} \quad (1)$$

In this table, the scalar products of different base vectors are equal to zero because of the mutual perpendicularity of the base vectors (see the property 3, Art. 163); $\mathbf{i}^2 = 1$, $\mathbf{j}^2 = 1$, $\mathbf{k}^2 = 1$ since $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are unit vectors (see the property 5, Art. 163).

Resolving the vectors \mathbf{a} and \mathbf{b} with respect to the basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$ yields, by Theorem 19 of Art. 157,

$$\left. \begin{array}{l} \mathbf{a} = X_1\mathbf{i} + Y_1\mathbf{j} + Z_1\mathbf{k}, \\ \mathbf{b} = X_2\mathbf{i} + Y_2\mathbf{j} + Z_2\mathbf{k}. \end{array} \right\} \quad (2)$$

In virtue of the algebraic properties of the scalar product established in Art. 162, we may compute \mathbf{ab} by multiplying, term by term, the right-hand members of relations (2):

$$\begin{aligned}\mathbf{ab} = & X_1X_2\mathbf{i}^2 + X_1Y_2\mathbf{ij} + X_1Z_2\mathbf{ik} + \\ & + Y_1X_2\mathbf{ji} + Y_1Y_2\mathbf{j}^2 + Y_1Z_2\mathbf{jk} + \\ & + Z_1X_2\mathbf{ki} + Z_1Y_2\mathbf{kj} + Z_1Z_2\mathbf{k}^2.\end{aligned}$$

Applying the base vector multiplication table (1), we find

$$\mathbf{ab} = X_1X_2 + Y_1Y_2 + Z_1Z_2,$$

as was to be shown.

The theorem just proved may be phrased as follows: *The scalar product of two vectors is equal to the sum of the products of their corresponding coordinates.*

165. We shall now state a number of important corollaries to Theorem 20.

Corollary 1. *A necessary and sufficient condition for the perpendicularity of vectors*

$$\mathbf{a} = \{X_1, Y_1, Z_1\} \quad \text{and} \quad \mathbf{b} = \{X_2, Y_2, Z_2\}$$

is given by the relation

$$X_1X_2 + Y_1Y_2 + Z_1Z_2 = 0. \quad (3)$$

For, by Art. 163, vectors \mathbf{a} and \mathbf{b} are mutually perpendicular if, and only if, $\mathbf{ab} = 0$. But, from Theorem 20, we have $\mathbf{ab} = X_1X_2 + Y_1Y_2 + Z_1Z_2$. Consequently, vectors \mathbf{a} and \mathbf{b} are perpendicular if, and only if, relation (3) holds true.

Corollary 2. *The angle φ between vectors*

$$\mathbf{a} = \{X_1, Y_1, Z_1\} \quad \text{and} \quad \mathbf{b} = \{X_2, Y_2, Z_2\}$$

is determined by the relation

$$\cos \varphi = \frac{X_1X_2 + Y_1Y_2 + Z_1Z_2}{\sqrt{X_1^2 + Y_1^2 + Z_1^2} \sqrt{X_2^2 + Y_2^2 + Z_2^2}}. \quad (4)$$

For, by the definition of the scalar product, $\mathbf{ab} = |\mathbf{a}||\mathbf{b}| \cos \varphi$; hence

$$\cos \varphi = \frac{\mathbf{ab}}{|\mathbf{a}| \cdot |\mathbf{b}|}. \quad (5)$$

But we have, by Theorem 20, $\mathbf{ab} = X_1X_2 + Y_1Y_2 + Z_1Z_2$ and, by Theorem 16 (Art. 139), $|\mathbf{a}| = \sqrt{X_1^2 + Y_1^2 + Z_1^2}$, $|\mathbf{b}| = \sqrt{X_2^2 + Y_2^2 + Z_2^2}$. Substitution of these expressions in (5) gives the required formula (4).

Corollary 3. *If an axis u makes angles α, β, γ with the coordinate axes, the projection of an arbitrary vector $s = \{X, Y, Z\}$ upon the axis u is determined by the relation*

$$\text{proj}_u s = X \cos \alpha + Y \cos \beta + Z \cos \gamma. \quad (6)$$

To prove this, suppose that the direction of the axis u is specified by a unit vector e . According to formula (2) of Art. 160, we have $es = |e| \text{proj}_e s$. Now observe that $|e| = 1$ and $\text{proj}_e s = \text{proj}_u s$. Hence, $\text{proj}_u s = es$. Since the vector e has the direction of the axis u , it follows that the vector e and the axis u make the same angles (namely, α, β, γ) with the coordinate axes. Hence, by Theorem 16 (Art. 139), we conclude that

$$\text{proj}_x e = |e| \cos \alpha, \quad \text{proj}_y e = |e| \cos \beta, \quad \text{proj}_z e = |e| \cos \gamma;$$

but $|e| = 1$, and so

$$e = \{\cos \alpha, \cos \beta, \cos \gamma\}.$$

We thus have $s = \{X, Y, Z\}$ and $e = \{\cos \alpha, \cos \beta, \cos \gamma\}$; by Theorem 20, we find: $\text{proj}_u s = es = X \cos \alpha + Y \cos \beta + Z \cos \gamma$, as was to be proved.

166. Example 1. Given three points $A(1, 1, 1)$, $B(2, 2, 1)$ and $C(2, 1, 2)$; find the angle $\varphi = \angle BAC$.

Solution. Using Theorem 15 (Art. 135), we find

$$\overline{AB} = \{1, 1, 0\}, \quad \overline{AC} = \{1, 0, 1\}.$$

Hence, by the second corollary to Theorem 20, we have

$$\cos \varphi = \frac{1 \cdot 1 + 1 \cdot 0 + 0 \cdot 1}{\sqrt{1^2 + 1^2 + 0^2} \sqrt{1^2 + 0^2 + 1^2}} = \frac{1}{\sqrt{2} \cdot \sqrt{2}} = \frac{1}{2}.$$

Consequently, $\varphi = 60^\circ$.

Example 2. Given the points $A(1, 1, 1)$ and $B(4, 5, -3)$; find the projection of the vector \overline{AB} on the axis u making equal acute angles with the coordinate axes.

Solution. Let $\cos \alpha, \cos \beta, \cos \gamma$ be the direction cosines of the axis u ; by the conditions of the problem, they are all equal and positive (since α, β, γ are equal acute angles). Now, by relation (4) of Art. 140,

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1.$$

Hence, from the foregoing,

$$\cos \alpha = \cos \beta = \cos \gamma = \frac{1}{\sqrt{3}}.$$

By Theorem 15 (Art. 135),

$$\overline{AB} = \{3, 4, -4\}.$$

All that remains now is to use the third corollary to Theorem 20, that is, to apply formula (6); doing this, we find

$$\text{proj}_u \overline{AB} = 3 \cdot \frac{1}{\sqrt{3}} + 4 \cdot \frac{1}{\sqrt{3}} - 4 \cdot \frac{1}{\sqrt{3}} = \sqrt{3}.$$

THE VECTOR AND TRIPLE SCALAR PRODUCTS OF VECTORS

§ 55. The Vector Product and Its Basic Properties

167. We shall now define a new operation on vectors, known as vector multiplication of vectors; it will be assumed that a rectangular cartesian system of coordinates has been attached to space.

The vector product of a vector \mathbf{a} by a vector \mathbf{b} is defined as the vector (denoted by the symbol $[\mathbf{ab}]$) which is determined by the following three conditions:

(1) *the modulus of the vector $[\mathbf{ab}]$ is equal to $|\mathbf{a}||\mathbf{b}| \sin \varphi$, where φ is the angle between the vectors \mathbf{a} and \mathbf{b} ;*

(2) *the vector $[\mathbf{ab}]$ is perpendicular to each of the vectors \mathbf{a} and \mathbf{b} ;*

(3) *the direction of the vector $[\mathbf{ab}]$ relative to the vectors \mathbf{a} and \mathbf{b} is the same as the direction of the coordinate axis Oz relative to the coordinate axes Ox and Oy . More precisely, if the three vectors \mathbf{a} , \mathbf{b} and $[\mathbf{ab}]$ are all drawn from a common initial point, then $[\mathbf{ab}]$ must be so directed that the shortest rotation of \mathbf{a} to \mathbf{b} will be (when viewed from the terminal point of $[\mathbf{ab}]$) in the same direction as the shortest rotation of the positive semi-axis Ox into the positive semi-axis Oy (when viewed from some point on the positive semi-axis Oz).*

For definiteness, we shall assume that, in the chosen coordinate system, the shortest rotation of the positive semi-axis Ox into the positive semi-axis Oy is seen from points on the positive semi-axis Oz to be *in the counterclockwise direction*. Such a system of coordinates is called *right-handed*. A right-handed system may also be characterised as follows: If the thumb of the right hand is extended in the direction of the axis Ox , while its forefinger is extended in the direction of the axis Oy , then its middle finger will point in the direction of the axis Oz of that system *).

The direction assigned to the vector product $[\mathbf{ab}]$ is in accordance with our choice of a right-handed system of coordinates;

*) A coordinate system is called left-handed if the axes Ox , Oy and Oz are directed analogous to the thumb, forefinger and middle finger of the left hand.

namely, if \mathbf{a} , \mathbf{b} and $[\mathbf{ab}]$ are drawn from the same initial point, then the vector $[\mathbf{ab}]$ must be so directed that the shortest rotation of the first factor, \mathbf{a} , to the second factor, \mathbf{b} , is seen from the terminal point of $[\mathbf{ab}]$ to be *in the counterclockwise direction* (Fig. 97). We may also apply the following "*rule of the right hand*": If \mathbf{a} , \mathbf{b} and $[\mathbf{ab}]$ are drawn from the same initial point,

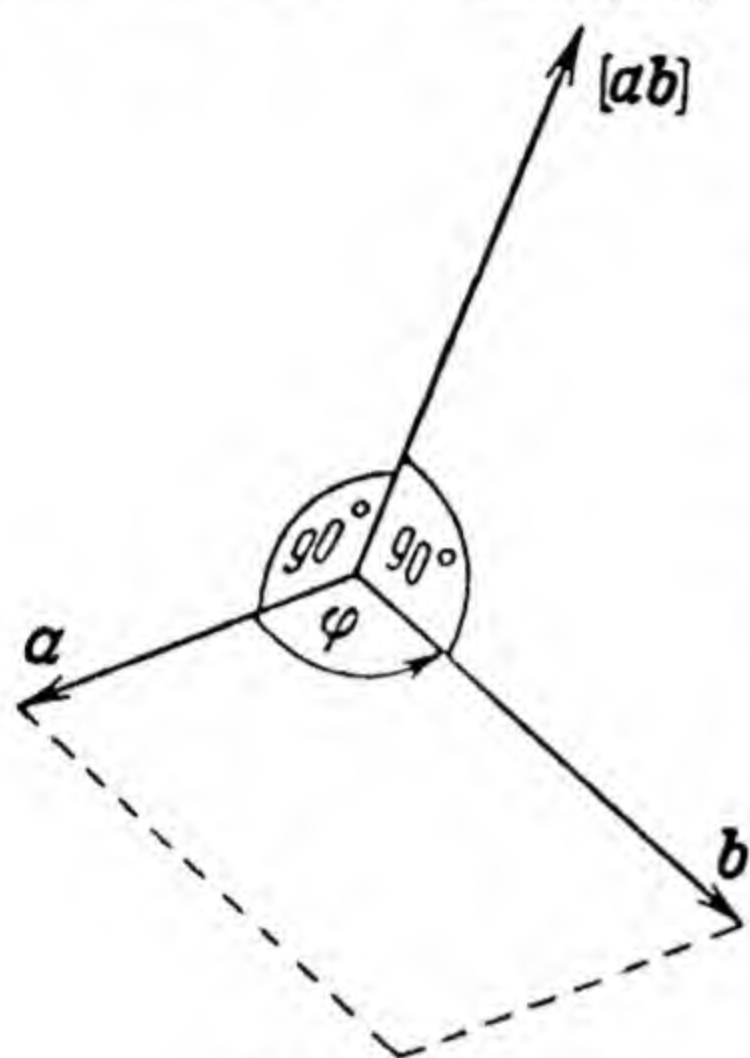


Fig. 97.

then the vector $[\mathbf{ab}]$ must be directed analogous to the middle finger of the right hand whose thumb extends in the direction of the first factor (that is, the vector \mathbf{a}), while its forefinger extends in the direction of the second factor (that is, the vector \mathbf{b}). It is this rule that will generally be referred to below.

168. The concept of the vector product owes its origin to mechanics; namely, if \mathbf{b} is the vector representing a force applied at a point M , and \mathbf{a} is the vector extending from a point O to the point M , then the vector $[\mathbf{ab}]$ represents the moment of the force \mathbf{b} about the point O .

In vector analysis $[\mathbf{ab}]$ is termed a *product* of vectors because it possesses some algebraic properties of the product of numbers (see the properties 2 and 3, Art. 171); it is termed the *vector product* because it is a vector.

169. First of all, we shall state the more important *geometric* properties of the vector product.

1. If vectors \mathbf{a} and \mathbf{b} are collinear, their vector product is equal to zero.

Proof. If \mathbf{a} and \mathbf{b} are collinear vectors, the angle φ between them is equal either to 0 (in the case when \mathbf{a} and \mathbf{b} are similarly directed) or to 180° (in the case when \mathbf{a} and \mathbf{b} are oppositely directed). In both cases, $\sin \varphi = 0$. Consequently, $|\mathbf{ab}| = |\mathbf{a}||\mathbf{b}| \sin \varphi = 0$, that is, the modulus of $[\mathbf{ab}]$ is zero, and hence the vector $[\mathbf{ab}]$ itself is equal to zero.

2. If the vector product of vectors \mathbf{a} and \mathbf{b} is equal to zero, the vectors \mathbf{a} and \mathbf{b} are collinear.

Proof. Let $[\mathbf{ab}] = 0$; then $|\mathbf{ab}| = 0$, and therefore $|\mathbf{a}||\mathbf{b}| \sin \varphi = 0$. If neither of the vectors \mathbf{a} and \mathbf{b} is equal to zero, then the last relation gives $\sin \varphi = 0$, and hence \mathbf{a} and \mathbf{b} are collinear vec-

tors. If, on the other hand, at least one of the vectors \mathbf{a} and \mathbf{b} is zero, then we may consider it collinear with the other vector, since a zero vector can be regarded as having any direction.

We now combine these two properties of the vector product into the following single statement: *Two vectors have their vector product zero if, and only if, they are collinear.*

3. *If vectors \mathbf{a} and \mathbf{b} are drawn from the same initial point, then the modulus of their vector product $[\mathbf{ab}]$ is equal to the area of the parallelogram constructed on the vectors \mathbf{a} and \mathbf{b} .*

Proof. Let S denote the area of the parallelogram constructed on the vectors \mathbf{a} and \mathbf{b} . It is known from elementary geometry that the area of a parallelogram is equal to the product of its adjacent sides by the sine of the included angle. Hence $|\mathbf{a}||\mathbf{b}|\sin\varphi = S$, and so

$$|[\mathbf{ab}]| = S, \quad (1)$$

as was to be shown.

170. In consequence of the last property, the vector product may be expressed by the formula

$$[\mathbf{ab}] = Se, \quad (2)$$

where \mathbf{e} is the vector determined by the following three conditions:

- (1) the modulus of \mathbf{e} is equal to unity;
- (2) the vector \mathbf{e} is perpendicular to each of the vectors \mathbf{a} and \mathbf{b} ;
- (3) the vector \mathbf{e} is directed in the same way as the middle finger of the right hand whose thumb extends in the direction of the vector \mathbf{a} , and whose forefinger extends in the direction of the vector \mathbf{b} (the vectors \mathbf{a} , \mathbf{b} and \mathbf{e} are assumed to be drawn from the same initial point).

To prove formula (2), compare the conditions which determine the vector \mathbf{e} with those determining the vector product $[\mathbf{ab}]$; from this comparison it will easily be seen that the vectors $[\mathbf{ab}]$ and \mathbf{e} are collinear and like-directed. Hence the vector $[\mathbf{ab}]$ can be obtained by multiplying the vector \mathbf{e} by a certain positive number; this number is equal to the ratio of the modulus of $[\mathbf{ab}]$ to the modulus of \mathbf{e} , and since $|\mathbf{e}| = 1$, it is equal simply to the modulus of $[\mathbf{ab}]$, i. e., to the number S . Thus, $[\mathbf{ab}] = Se$, as was to be proved. (Fig. 98 is an illustration of formula (2) for the case $|\mathbf{a}| = 2$, $|\mathbf{b}| = 2$, $\varphi = 90^\circ$.)

171. We now proceed to establish the algebraic properties of the vector product.

1. The vector product is anticommutative:

$$[ab] = -[ba]; \quad (3)$$

that is, the vector product of a by b is the negative of the vector product of b by a .

Proof. If a and b are collinear vectors, then both $[ab]$ and $[ba]$ are zero, so that relation (3) holds. Suppose now that the vectors a and b are not collinear.

To begin with, note that, by the first two conditions contained in the definition of the vector product, the vectors $[ab]$ and $[ba]$ have the same modulus and are collinear; hence, either $[ab] = [ba]$ or $[ab] = -[ba]$. It remains to determine which of these two possibilities materialises. The question is resolved by a considera-

tion of the third condition. For, when extending the thumb and forefinger of the right hand first in the directions of a and b , respectively, and then in the directions of b and a , the hand has to be turned in such a way that the direction of the middle finger in the second case will be opposite to its direction in the first case. Hence $[ab]$ and $[ba]$ have opposite directions, that is, $[ab] = -[ba]$.

2. The vector product is associative with respect to multiplication by a scalar:

$$[(\lambda a)b] = \lambda [ab] \quad (4)$$

and

$$[a(\lambda b)] = \lambda [ab]. \quad (5)$$

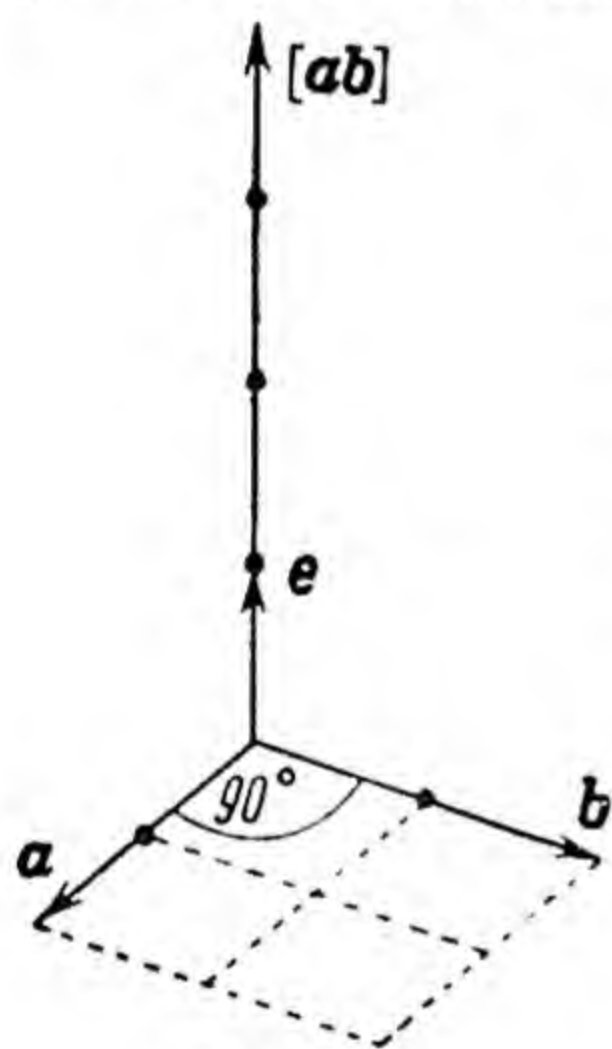


Fig. 98.

Proof. Formula (5) can be reduced to (4) by reversing the order of the factors of the vector products in both the left and the right member (whereupon the letters a and b must be interchanged). Hence, it will be sufficient to establish formula (4).

Note first of all that, if $\lambda = 0$ or if a, b are collinear vectors, formula (4) is obviously true since, in either case, its left and right members will be equal to zero. Suppose now that $\lambda \neq 0$ and the vectors a, b are not collinear.

By the first condition in the definition of the vector product, the modulus of the vector $[ab]$ is equal to $|a||b|\sin\varphi$, where φ is the angle between the vectors a and b ; hence the modulus of the vector $\lambda[ab]$ is equal to $|\lambda||a||b|\sin\varphi$. By the same condition, the modulus of the vector $[(\lambda a)b]$ is equal to $|\lambda||a||b|\sin\psi$,

where ψ is the angle between the vectors $\lambda \mathbf{a}$ and \mathbf{b} . But the angle ψ is equal either to the angle φ (if λ is positive), or to the angle $\pi - \varphi$ (if λ is negative); in both cases, $\sin \varphi = \sin \psi$. It follows that the modulus of the vector $[(\lambda \mathbf{a}) \mathbf{b}]$ is equal to the modulus of the vector $\lambda [\mathbf{a} \mathbf{b}]$.

By the second condition contained in the definition of the vector product, both vectors $\lambda [\mathbf{a} \mathbf{b}]$ and $[(\lambda \mathbf{a}) \mathbf{b}]$ are perpendicular to each of the vectors \mathbf{a} and \mathbf{b} ; the vectors $\lambda [\mathbf{a} \mathbf{b}]$ and $[(\lambda \mathbf{a}) \mathbf{b}]$ are therefore collinear.

Since the vectors $\lambda [\mathbf{a} \mathbf{b}]$ and $[(\lambda \mathbf{a}) \mathbf{b}]$ have the same modulus and are collinear, it follows that they are equal or else are the negatives of each other; that is, either $[(\lambda \mathbf{a}) \mathbf{b}] = \lambda [\mathbf{a} \mathbf{b}]$, or $[(\lambda \mathbf{a}) \mathbf{b}] = -\lambda [\mathbf{a} \mathbf{b}]$. It remains to find out which of the two possibilities materialises. We shall have to consider the cases $\lambda > 0$ and $\lambda < 0$ separately.

Let $\lambda > 0$; then the vectors $\lambda \mathbf{a}$ and \mathbf{a} have the same direction. In this case, by the rule of the right hand, the vector $[(\lambda \mathbf{a}) \mathbf{b}]$ has the same direction as the vector $[\mathbf{a} \mathbf{b}]$; but, if $\lambda > 0$, the vector $\lambda [\mathbf{a} \mathbf{b}]$ also has the same direction as $[\mathbf{a} \mathbf{b}]$. Consequently, the vectors $[(\lambda \mathbf{a}) \mathbf{b}]$ and $\lambda [\mathbf{a} \mathbf{b}]$ are like-directed, and hence $[(\lambda \mathbf{a}) \mathbf{b}] = \lambda [\mathbf{a} \mathbf{b}]$. Now let $\lambda < 0$; then the vectors $\lambda \mathbf{a}$ and \mathbf{a} have opposite directions. In this case, by the rule of the right hand, the vector $[(\lambda \mathbf{a}) \mathbf{b}]$ has its direction opposite to that of $[\mathbf{a} \mathbf{b}]$; but, if $\lambda < 0$, the vector $\lambda [\mathbf{a} \mathbf{b}]$ also has the opposite direction to that of $[\mathbf{a} \mathbf{b}]$. Consequently, the vectors $[(\lambda \mathbf{a}) \mathbf{b}]$ and $\lambda [\mathbf{a} \mathbf{b}]$ are like-directed, and hence $[(\lambda \mathbf{a}) \mathbf{b}] = \lambda [\mathbf{a} \mathbf{b}]$. We thus see that this relation is always valid.

3. The vector product is distributive with respect to addition:

$$[\mathbf{a} (\mathbf{b} + \mathbf{c})] = [\mathbf{a} \mathbf{b}] + [\mathbf{a} \mathbf{c}] \quad (6)$$

and

$$[(\mathbf{b} + \mathbf{c}) \mathbf{a}] = [\mathbf{b} \mathbf{a}] + [\mathbf{c} \mathbf{a}]. \quad (7)$$

Proof. Formula (7) can be reduced to (6) by reversing the order of the factors in both the left and the right member of (7). It will, therefore, be sufficient to establish formula (6). It should also be noted that formula (6) is clearly true if $\mathbf{a} = 0$. We shall assume in what follows that $\mathbf{a} \neq 0$.

Let us begin by considering the particular case when the first of the vectors is a unit vector and the other two vectors are each perpendicular to the first vector.

Draw all the three vectors from a common initial point O . Let \mathbf{a}_0 be the first (unit) vector; denote the other two vectors (perpendicular to \mathbf{a}_0) by \overline{OB} and \overline{OC} , and their sum by \overline{OD} , i. e., $\overline{OD} = \overline{OB} + \overline{OC}$ (Fig. 99).

Introducing the notation

$$\overline{OB}^* = [\mathbf{a}_0 \overline{OB}], \quad \overline{OC}^* = [\mathbf{a}_0 \overline{OC}], \quad \overline{OD}^* = [\mathbf{a}_0 \overline{OD}] = [\mathbf{a}_0 (\overline{OB} + \overline{OC})],$$

we have, by virtue of the first two conditions in the definition of the vector product,

$$(1) \quad |\overline{OB}^*| = |[\mathbf{a}_0 \overline{OB}]| = |\mathbf{a}_0| |\overline{OB}| \sin 90^\circ = |\overline{OB}|.$$

$$(2) \quad \overline{OB}^* \perp \mathbf{a}_0, \quad \overline{OB}^* \perp \overline{OB}.$$

It follows that the vector \overline{OB}^* can be obtained by rotating the vector \overline{OB} about \mathbf{a}_0 through 90° . In virtue of the third condition, this rotation will be in the counterclockwise direction (when

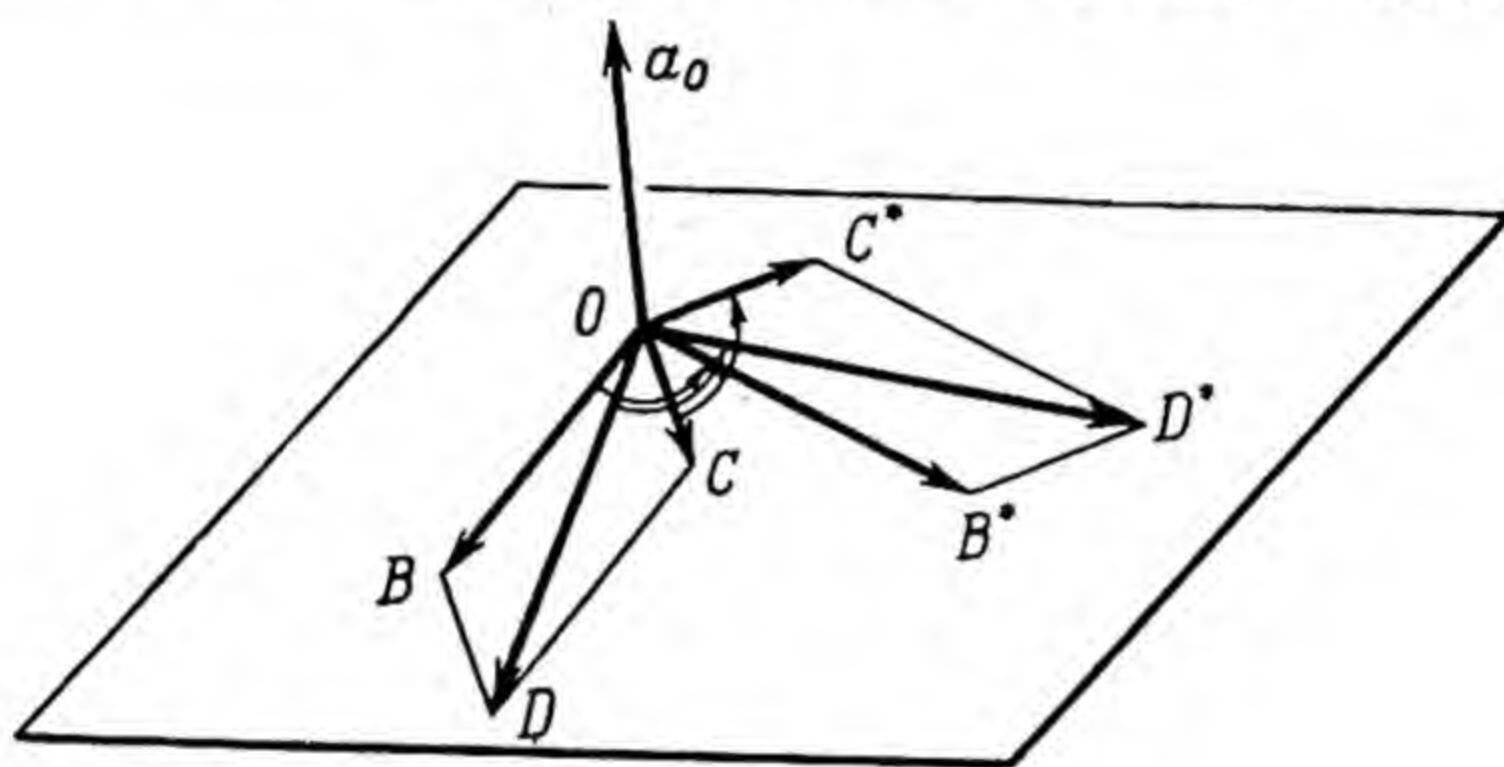


Fig. 99.

viewed from the terminal point of the vector \mathbf{a}_0). In like manner, the vectors \overline{OC}^* and \overline{OD}^* are obtained by rotating the vectors \overline{OC} and \overline{OD} about \mathbf{a}_0 through 90° and in the same direction. Thus, the entire figure $OB^*D^*C^*$ is obtained by a rotation of the parallelogram $OBDC$; consequently, $OB^*D^*C^*$ is a parallelogram. Hence we conclude that $\overline{OD}^* = \overline{OB}^* + \overline{OC}^*$, or

$$[\mathbf{a}_0 \overline{OD}] = [\mathbf{a}_0 \overline{OB}] + [\mathbf{a}_0 \overline{OC}]. \quad (8)$$

This is relation (6) for our particular case.

Next, let \mathbf{a} be any vector perpendicular to the vectors \overline{OB} and \overline{OC} . Let \mathbf{a}_0 denote the unit vector having the same direction as \mathbf{a} ; then $\mathbf{a} = |\mathbf{a}| \mathbf{a}_0$. Multiplying both members of (8) by the number $|\mathbf{a}|$ and replacing $|\mathbf{a}| \mathbf{a}_0$ by \mathbf{a} , we obtain

$$[\mathbf{a} \overline{OD}] = [\mathbf{a} \overline{OB}] + [\mathbf{a} \overline{OC}]. \quad (9)$$

Finally, consider vectors \mathbf{a} , \mathbf{b} , \mathbf{c} having arbitrary directions. Let us assume that they are drawn from a common initial point O .

Through the terminal points of the vectors \mathbf{b} , \mathbf{c} and $\mathbf{b} + \mathbf{c}$, draw straight lines parallel to the vector \mathbf{a} . Pass a plane through the point O perpendicular to these lines and intersecting them in the points B , C and D , respectively (Fig. 100).

Consider the vector products $[\mathbf{ab}]$ and $[\mathbf{a} \overline{OB}]$; they are easily shown to give the same vector. For, in the first place, the modulus of $[\mathbf{ab}]$ is equal to the modulus of $[\mathbf{a} \overline{OB}]$, since the area of the parallelogram constructed on the vectors \mathbf{a} and \mathbf{b} is equal to the area of the rectangle constructed on the vectors \mathbf{a} and \overline{OB} ; secondly, the vectors $[\mathbf{ab}]$ and $[\mathbf{a} \overline{OB}]$ are collinear, since both are perpendicular to the same plane (namely, to the plane containing the vectors \mathbf{a} , \mathbf{b} and \overline{OB}); and lastly, in accordance with the rule of the right hand, the vectors $[\mathbf{ab}]$ and $[\mathbf{a} \overline{OB}]$ are like-directed. Thus,

$$[\mathbf{a} \overline{OB}] = [\mathbf{ab}].$$

Similarly,

$$[\mathbf{a} \overline{OC}] = [\mathbf{ac}], \quad [\mathbf{a} \overline{OD}] = [\mathbf{a}(\mathbf{b} + \mathbf{c})].$$

Inserting these expressions in (9), we obtain

$$[\mathbf{a}(\mathbf{b} + \mathbf{c})] = [\mathbf{ab}] + [\mathbf{ac}],$$

as was to be proved.

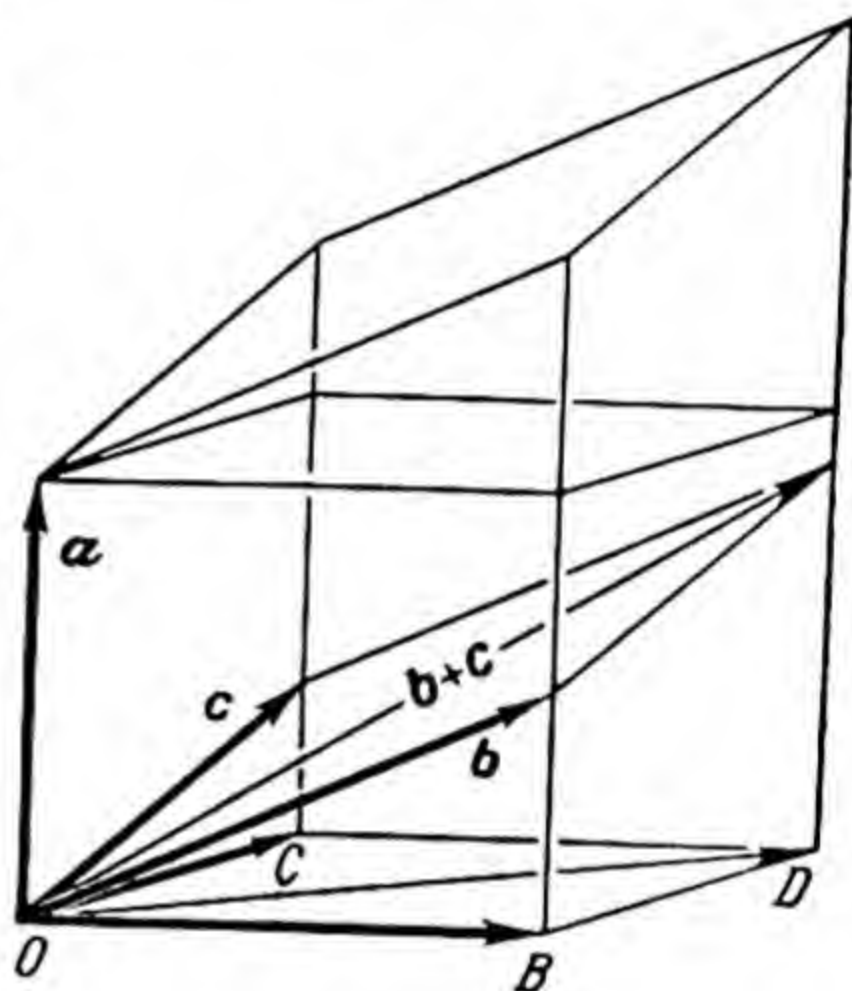


Fig. 100.

172. The last of the algebraic properties just established enables us to carry out the vector multiplication of vector polynomials term by term. The second property of the vector product permits us to group together the scalar coefficients of the vector factors. For example,

$$\begin{aligned} [(2\mathbf{a} + 5\mathbf{b})(3\mathbf{c} + 4\mathbf{d})] &= [(2\mathbf{a} + 5\mathbf{b})(3\mathbf{c})] + [(2\mathbf{a} + 5\mathbf{b})(4\mathbf{d})] = \\ &= [(2\mathbf{a})(3\mathbf{c})] + [(5\mathbf{b})(3\mathbf{c})] + [(2\mathbf{a})(4\mathbf{d})] + [(5\mathbf{b})(4\mathbf{d})] = \\ &= 6[\mathbf{ac}] + 15[\mathbf{bc}] + 8[\mathbf{ad}] + 20[\mathbf{bd}]. \end{aligned}$$

It must, however, be borne in mind that the *order of the factors* of the vector product is material. In consequence of the first property (Art. 171), a *minus sign* must be prefixed to the vector product after its factors have been interchanged.

Note. As has been shown in Art. 169, collinear vectors have their vector product zero. In particular, the vector product of two identical factors is zero: $[aa] = 0$. For this reason, the concept of the vector square is not used in vector analysis.

§ 56. Representation of the Vector Product in Terms of the Coordinates of the Vector Factors

173. The following theorem permits us to compute the vector product of two vectors from their coordinates, that is, from their projections on the axes of a rectangular cartesian coordinate system.

Theorem 21. *Given the coordinates of vectors a and b :*

$$a = \{X_1, Y_1, Z_1\}, \quad b = \{X_2, Y_2, Z_2\},$$

the vector product of the vector a by the vector b is determined by the formula

$$[ab] = \left\{ \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}, \quad - \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}, \quad \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \right\}. \quad (1)$$

Proof. We shall first draw up the vector multiplication table for the base vectors. According to the note made at the close of Art. 172, $[ii] = 0$, $[jj] = 0$, $[kk] = 0$. Consider now the vector product $[ij]$. The modulus of the vector $[ij]$ is equal to the area of the parallelogram constructed on the vectors i and j (see Art. 169, the property 3). This parallelogram is a square with side equal to unity, so that its area is equal to unity. Hence $[ij]$ is a unit vector. Since the vector $[ij]$ must be perpendicular to the vectors i and j and directed according to the rule of the right hand, it is easily seen that $[ij]$ coincides with the third base vector k ; that is, $[ij] = k$. By a similar argument, we find the relations $[jk] = i$, $[ki] = j$. It remains to express $[ji]$, $[kj]$, and $[ik]$; but $[ji] = -[ij]$, $[kj] = -[jk]$, $[ik] = -[ki]$; hence $[ji] = -k$, $[kj] = -i$, $[ik] = -j$. Thus, the desired multiplication table is as follows:

$$\left. \begin{array}{lll} [ii] = 0, & [ij] = k, & [ik] = -j, \\ [ji] = -k, & [jj] = 0, & [jk] = i, \\ [ki] = j, & [kj] = -i, & [kk] = 0. \end{array} \right\} \quad (2)$$

By Theorem 19 of Art. 157, the resolution of the vectors \mathbf{a} and \mathbf{b} with respect to the basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$ yields

$$\left. \begin{aligned} \mathbf{a} &= X_1\mathbf{i} + Y_1\mathbf{j} + Z_1\mathbf{k}, \\ \mathbf{b} &= X_2\mathbf{i} + Y_2\mathbf{j} + Z_2\mathbf{k}. \end{aligned} \right\} \quad (3)$$

In virtue of the algebraic properties of the vector product, established in Art. 171, we may compute $[\mathbf{ab}]$ by multiplying, term by term, the right-hand members of relations (3):

$$\begin{aligned} [\mathbf{ab}] &= X_1X_2[\mathbf{ii}] + X_1Y_2[\mathbf{ij}] + X_1Z_2[\mathbf{ik}] + \\ &\quad + Y_1X_2[\mathbf{ji}] + Y_1Y_2[\mathbf{jj}] + Y_1Z_2[\mathbf{jk}] + \\ &\quad + Z_1X_2[\mathbf{ki}] + Z_1Y_2[\mathbf{kj}] + Z_1Z_2[\mathbf{kk}]. \end{aligned}$$

Making use of the base vector multiplication table (2), we hence find

$$[\mathbf{ab}] = (Y_1Z_2 - Y_2Z_1)\mathbf{i} - (X_1Z_2 - X_2Z_1)\mathbf{j} + (X_1Y_2 - X_2Y_1)\mathbf{k},$$

or

$$[\mathbf{ab}] = \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix} \mathbf{i} - \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix} \mathbf{j} + \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \mathbf{k}. \quad (4)$$

We have expressed the vector $[\mathbf{ab}]$ in terms of the basis $\mathbf{i}, \mathbf{j}, \mathbf{k}$; the coefficients of this resolution are the coordinates of the vector $[\mathbf{ab}]$. Thus,

$$[\mathbf{ab}] = \left\{ \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}, -\begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}, \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \right\}, \quad (1)$$

as was to be shown.

Note. When carrying out calculations by formula (1), it is convenient to begin by forming the following array from the coordinates of the given vectors:

$$\begin{pmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{pmatrix}.$$

Covering in turn the first, the second and, finally, the third column of this array gives three determinants of the second order; evaluating the determinants and prefixing a minus sign to the second of them, we find the three coordinates of the vector product $[\mathbf{ab}]$.

It should also be remarked that formula (4), which is equivalent to (1), can be put in the form

$$[ab] = \begin{vmatrix} i & j & k \\ X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \end{vmatrix}. \quad (5)$$

For, on expanding this determinant in terms of the elements of the first row, we get an expression identical with the right-hand member of (4).

174. Example 1. Given the vectors $\mathbf{a} = \{2, 5, 7\}$ and $\mathbf{b} = \{1, 2, 4\}$. Find the coordinates of the vector product $[ab]$.

Solution. In accordance with the note made at the end of the preceding article, we form the array

$$\begin{pmatrix} 2 & 5 & 7 \\ 1 & 2 & 4 \end{pmatrix}.$$

By covering in turn the columns of this array, we obtain three determinants of the second order; evaluating them and taking the second determinant with a minus sign, we find the required projections

$$[ab] = \{6, -1, -1\}.$$

Example 2. $A(1, 1, 1)$, $B(2, 2, 2)$ and $C(4, 3, 5)$ are three given points in space. Find the area S_Δ of the triangle ABC .

Solution. Consider the vectors \overline{AB} and \overline{AC} . By the property 3 (Art. 169), the modulus of the vector product $[\overline{AB} \overline{AC}]$ is equal to the area of the parallelogram constructed on the vectors \overline{AB} and \overline{AC} . Now, the required area S_Δ of the triangle ABC is equal to half the area of this parallelogram; hence

$$S_\Delta = \frac{1}{2} |[\overline{AB} \overline{AC}]|.$$

It only remains to compute the right-hand member of this relation.

We first find, by Theorem 15 (Art. 135), the coordinates of the vectors \overline{AB} and \overline{AC} :

$$\overline{AB} = \{1, 1, 1\}, \quad \overline{AC} = \{3, 2, 4\}.$$

$$\text{Hence, } [\overline{AB} \overline{AC}] = \{2, -1, -1\} \text{ and } |[\overline{AB} \overline{AC}]| = \sqrt{2^2 + 1^2 + 1^2} = \sqrt{6}.$$

$$\text{Thus, } S_\Delta = \frac{1}{2} \sqrt{6}.$$

§ 57. The Triple Scalar Product

175. Let there be given any three vectors \mathbf{a} , \mathbf{b} and \mathbf{c} . Suppose that the *vector* multiplication of the vectors \mathbf{a} and \mathbf{b} is followed by the *scalar* multiplication of the resulting vector $[ab]$ and the vector \mathbf{c} ; the *number* $[ab]\mathbf{c}$ thus determined is called the *triple scalar product* of the three vectors \mathbf{a} , \mathbf{b} , \mathbf{c} . In the remaining arti-

cles of this chapter, we shall study the basic properties of the triple scalar product and point out some problems where the triple scalar product can effectively be used.

176. We shall agree to speak of vectors \mathbf{a} , \mathbf{b} , \mathbf{c} as *coplanar* if they lie in the same plane or in parallel planes. Since geometric vectors are free vectors, it follows that coplanar vectors can always be made to lie in the same plane by means of parallel displacement (translation). In particular, coplanar vectors will lie in the same plane when they are drawn from a common initial point.

177. If three given vectors are also designated as the first, second and third vector, then the given vectors are called an *ordered triad* of vectors; we shall hereafter refer to them simply as a *triad* of vectors, omitting the adjective. In the text, a triad of vectors will be written in order; for example, when we write \mathbf{a} , \mathbf{b} , \mathbf{c} , this will mean that \mathbf{a} is regarded as the first, \mathbf{b} as the second, and \mathbf{c} as the third vector, whereas writing \mathbf{b} , \mathbf{c} , \mathbf{a} , will mean that \mathbf{b} is regarded as the first, \mathbf{c} as the second, and \mathbf{a} as the third vector.

178. A triad of non-coplanar vectors is called *right-handed* if its vectors, when drawn from a common initial point and taken in their order, are directed analogous to the thumb, forefinger and middle finger of the right hand. Speaking in more detail, a triad of non-coplanar vectors is called a right-handed triad if its third vector is on the same side of the plane containing the first two vectors, as the middle finger of the right hand whose thumb extends in the direction of the first vector and whose forefinger extends in the direction of the second vector of the triad.

A triad of non-coplanar vectors is called *left-handed* if its vectors, when drawn from a common initial point and taken in their order, are directed analogous to the thumb, forefinger and middle finger of the left hand.

Triads of coplanar vectors are neither right-handed nor left-handed.

179. Let \mathbf{a} , \mathbf{b} , \mathbf{c} be any given non-coplanar vectors. Numbering them in all possible different ways, we obtain six triads: \mathbf{a} , \mathbf{b} , \mathbf{c} ; \mathbf{b} , \mathbf{c} , \mathbf{a} ; \mathbf{c} , \mathbf{a} , \mathbf{b} ; \mathbf{b} , \mathbf{a} , \mathbf{c} ; \mathbf{a} , \mathbf{c} , \mathbf{b} ; \mathbf{c} , \mathbf{b} , \mathbf{a} . By the inspection of a model (which is easily made of wire), it can be verified that three of these triads are right-handed, and the other three left-handed; namely,

$$\mathbf{a}, \mathbf{b}, \mathbf{c}; \quad \mathbf{b}, \mathbf{c}, \mathbf{a}; \quad \mathbf{c}, \mathbf{a}, \mathbf{b}$$

are triads of one orientation, that is, are either all of them right-handed, or all left-handed;

$$b, a, c; \quad a, c, b; \quad c, b, a$$

are triads of the opposite orientation *).

180. The *geometric* meaning of the triple scalar product is expressed by the following important

Theorem 22. *The triple scalar product $[ab]c$ is equal to the volume of the parallelepiped constructed on the vectors a, b, c ; the sign of this volume is positive or negative according as the*

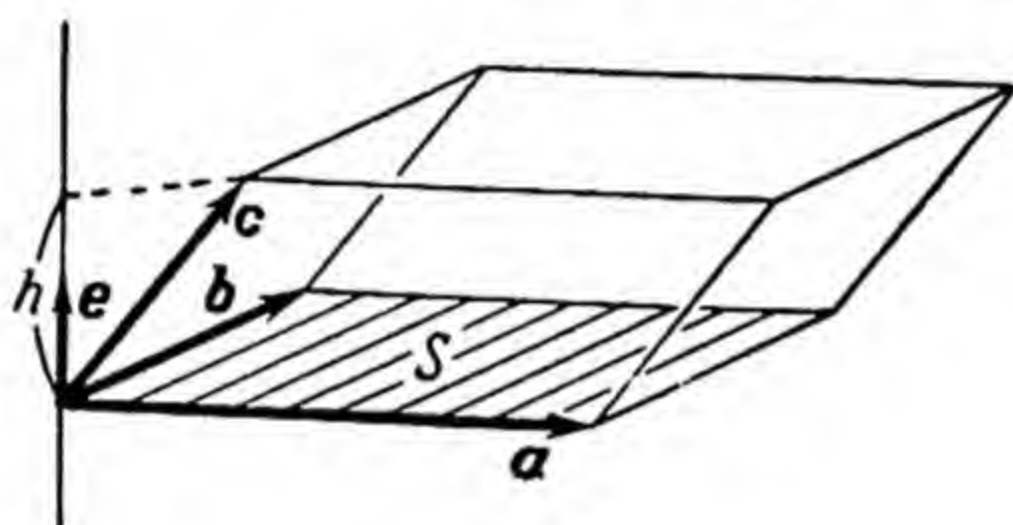


Fig. 101.

triad a, b, c is right-handed or left-handed. If the vectors a, b, c are coplanar, $[ab]c = 0$.

Proof. Suppose first that the vectors a and b are *non-collinear*. Denote by S the area of the parallelogram constructed on the vectors a, b , and let e be a unit vector, defined as in Art. 170. By formula (2) of Art. 170, we have

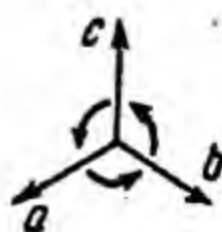
$$[ab] = Se.$$

Hence

$$[ab]c = S(ec) = S|e| \text{proj}_e c = S \text{proj}_e c. \quad (1)$$

But $\text{proj}_e c = \pm h$, where h is the altitude of the parallelepiped constructed on the vectors a, b, c and having as its base the

*) The following visual method for distinguishing triads can also be used. Imagine that you are placed inside the solid angle made by a given triad of vectors. If the circuit from the first to the second vector, from the second to the third and, finally, from the third to the first vector is then seen to be in the *counterclockwise* direction, the given triad is *right-handed*; if this circuit is seen to be in the *clockwise* direction, the triad is *left-handed*. According to this rule, the triads a, b, c ; b, c, a and c, a, b in the diagram given here are right-handed,



parallelogram constructed on the vectors \mathbf{a} , \mathbf{b} (Fig. 101). Hence, denoting the volume of the parallelepiped by V and recalling that $hS = V$, we find from (1):

$$[\mathbf{ab}]c = \pm V. \quad (2)$$

We must now determine when the sign of the volume V will be positive, and when negative. For this purpose, note that $\text{proj}_e c = +h$ if the vector c is on the same side of the base plane (containing \mathbf{a} , \mathbf{b}) as the vector e , that is, if the triad \mathbf{a} , \mathbf{b} , \mathbf{c} is of the same orientation as the triad \mathbf{a} , \mathbf{b} , \mathbf{e} (see Art. 178); $\text{proj}_e c = -h$ if the vectors c and e are on opposite sides of the plane containing \mathbf{a} , \mathbf{b} , that is, if the triads \mathbf{a} , \mathbf{b} , \mathbf{c} and \mathbf{a} , \mathbf{b} , \mathbf{e} have opposite orientations. Now, by the definition of the vector e (see Art. 170), the triad \mathbf{a} , \mathbf{b} , \mathbf{e} is right-handed. Hence, the volume V in (2) is positive if \mathbf{a} , \mathbf{b} , \mathbf{c} form a right-handed triad, and negative if \mathbf{a} , \mathbf{b} , \mathbf{c} form a left-handed triad. If the vector c lies in the plane of \mathbf{a} , \mathbf{b} , that is, if the vectors \mathbf{a} , \mathbf{b} , \mathbf{c} are coplanar, then $\text{proj}_e c = 0$ and, as is clear from (1), $[\mathbf{ab}]c = 0$. This completes the proof of the theorem for the case when the vectors \mathbf{a} , \mathbf{b} are non-collinear. There remains for consideration the case when \mathbf{a} , \mathbf{b} are collinear. In this case, $[\mathbf{ab}] = 0$, and hence $[\mathbf{ab}]c = 0$, which again is in accordance with what has been stated in the theorem (since, if the vectors \mathbf{a} and \mathbf{b} are collinear, the three vectors \mathbf{a} , \mathbf{b} , \mathbf{c} are coplanar).

The theorem is thus proved.

181. The following identity is easily derived from Theorem 22:

$$[\mathbf{ab}]c = a[\mathbf{bc}]. \quad (3)$$

Proof. Since the scalar product is commutative,

$$a[\mathbf{bc}] = [\mathbf{bc}]a. \quad (4)$$

Further, by Theorem 22 we have

$$[\mathbf{ab}]c = \pm V, [\mathbf{bc}]a = \pm V. \quad (5)$$

According to Art. 179 the triads \mathbf{a} , \mathbf{b} , \mathbf{c} and \mathbf{b} , \mathbf{c} , \mathbf{a} are of one orientation; hence, by Theorem 22, the right members of relations (5) have like signs. Relations (5) thus yield

$$[\mathbf{ab}]c = [\mathbf{bc}]a.$$

Hence, in consequence of (4),

$$[\mathbf{ab}]c = a[\mathbf{bc}],$$

as was to be shown.

182. From now on, we shall denote the triple scalar products $[ab]c$ and $a[bc]$ by a simpler symbol: abc . No ambiguity can arise from the omission of the brackets indicating vector multiplication, since $[ab]c = a[bc]$.

183. It should be stressed that Theorem 22 has as its immediate consequence the following proposition:

The triple scalar product of vectors a, b, c is zero if, and only if, the vectors a, b, c are coplanar.

For, Theorem 22 plainly states that $abc = 0$ if a, b, c are coplanar vectors. The fact that $abc = 0$ only if a, b, c are coplanar vectors follows from the same theorem; for, if the vectors a, b, c are non-coplanar, then the parallelepiped constructed on them has a volume different from zero, and hence $abc \neq 0$.

The same proposition can also be phrased as follows:

A necessary and sufficient condition for the coplanarity of three vectors a, b, c is that their triple scalar product should be zero: $abc = 0$.

§ 58. Representation of the Triple Scalar Product in Terms of the Coordinates of the Vector Factors

184. Theorem 23. *Given the coordinates of vectors a, b, c :*

$$a = \{X_1, Y_1, Z_1\}, \quad b = \{X_2, Y_2, Z_2\}, \quad c = \{X_3, Y_3, Z_3\},$$

the triple scalar product abc is determined by the formula

$$abc = \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}.$$

Proof. We have $abc = [ab]c$. From Theorem 21 (Art. 173),

$$[ab] = \left\{ \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix}, -\begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix}, \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} \right\}.$$

The scalar multiplication of this vector by the vector $c = \{X_3, Y_3, Z_3\}$ yields, by Theorem 20 (Art. 164),

$$\begin{aligned} abc = [ab]c &= \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix} X_3 - \begin{vmatrix} X_1 & Z_1 \\ X_2 & Z_2 \end{vmatrix} Y_3 + \begin{vmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{vmatrix} Z_3 = \\ &= \begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix}, \end{aligned}$$

as was to be proved.

Example. $A(1, 1, 1)$, $B(4, 4, 4)$, $C(3, 5, 5)$, $D(2, 4, 7)$ are four given points in space. Find the volume of the tetrahedron $ABCD$.

Solution. From elementary geometry, the volume V_T of the tetrahedron $ABCD$ is one sixth of the volume of the parallelepiped constructed on the vectors \overline{AB} , \overline{AC} and \overline{AD} ; hence, from Theorem 22, we conclude that V_T is numerically equal to one sixth of the triple scalar product $\overline{AB} \cdot \overline{AC} \cdot \overline{AD}$. It only remains to compute the value of this triple product. First of all, we determine the coordinates of the vectors \overline{AB} , \overline{AC} , \overline{AD} . By Theorem 15 (Art. 135), we have: $\overline{AB} = \{3, 3, 3\}$, $\overline{AC} = \{2, 4, 4\}$, $\overline{AD} = \{1, 3, 6\}$.

Using Theorem 23, we now find

$$\overline{AB} \cdot \overline{AC} \cdot \overline{AD} = \begin{vmatrix} 3 & 3 & 3 \\ 2 & 4 & 4 \\ 1 & 3 & 6 \end{vmatrix} = 18.$$

Hence, $V_T = 3$.

185. According to Art. 183, *a necessary and sufficient condition for the coplanarity of three vectors is that their triple scalar product should be zero.*

Hence, by Theorem 23, we conclude: *Given the coordinates of vectors \mathbf{a} , \mathbf{b} , \mathbf{c} :*

$$\mathbf{a} = \{X_1, Y_1, Z_1\}, \quad \mathbf{b} = \{X_2, Y_2, Z_2\}, \quad \mathbf{c} = \{X_3, Y_3, Z_3\},$$

a necessary and sufficient condition for the coplanarity of these vectors is that

$$\begin{vmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y_2 & Z_2 \\ X_3 & Y_3 & Z_3 \end{vmatrix} = 0,$$

i. e., that the determinant of the third order formed from the coordinates of the vectors \mathbf{a} , \mathbf{b} , \mathbf{c} should be zero.

THE EQUATION OF A SURFACE AND THE EQUATIONS OF A CURVE

§ 59. The Equation of a Surface

186. As we know, some of the simpler surfaces (the plane, sphere, circular cylinder, circular cone) yield readily to investigation by the methods of elementary geometry. But the general problem of investigating the diverse curves encountered when dealing with various questions of mathematics and its applications demands the use of more advanced methods, furnished by algebra and mathematical analysis. This use of algebraic and analytical methods is based on a uniform mode of determining a surface, namely, that of representing a surface by an equation.

187. Let x, y, z be arbitrary variables; this means that the symbols x, y, z represent any (real) numbers whatsoever. A relation of the form $F(x, y, z) = 0$, where $F(x, y, z)$ denotes an expression containing x, y, z , is called an equation in three variables x, y, z (provided that $F(x, y, z) = 0$ is valid not identically, that is, not for every triad of numbers x, y, z).

Three numbers $x = x_0, y = y_0, z = z_0$ are said to *satisfy* a given equation in three variables if the equation holds true when these numbers are substituted in it for the variables. If $F(x, y, z) = 0$ is an identity, it is satisfied by any numbers x, y, z .

188. The fundamental concept of solid analytic geometry is that of the *equation of a surface*. We shall now explain the meaning of this concept.

Let there be given any surface in space; also, let a coordinate system be chosen.

The equation of a given surface (in a chosen coordinate system) is defined as the equation in three variables which is satisfied by the coordinates of all points lying on the surface and by the coordinates of no other point.

Thus, if the equation of a surface is known, we can determine for each point in space whether or not it lies on the surface. To answer this question, it is necessary merely to substitute the coordinates of the point for the variables in the equation; if the coordinates of the point under test satisfy the equation, the point

lies on the surface; whereas, if its coordinates do not satisfy the equation, the point does not lie on the surface.

The definition just made constitutes the basis of the methods of solid analytic geometry, which consist essentially in the investigation of surfaces by analysing their equations. In cases when the surface under consideration is determined in purely geometric terms, we begin its investigation by deriving the equation of the surface. In many problems, however, the equation of a surface is regarded as something known, while the surface itself is regarded as something to be derived. In other words, often an equation is given beforehand, and a surface is thereby determined.

189. If an equation is given and we are to answer the question: "What surface is represented by this equation?" (or: "What is the surface having this as its equation?"), then it is convenient to use the definition phrased as follows:

The surface represented by a given equation (referred to some coordinate system) is the locus of those points whose coordinates satisfy the equation.

Note. If $M(x, y, z)$ is a variable point of the surface, then x, y, z are called *the current coordinates*.

190. Example. In rectangular cartesian coordinates, the equation

$$(x - \alpha)^2 + (y - \beta)^2 + (z - \gamma)^2 = r^2 \quad (1)$$

represents a sphere, whose centre is at the point $C(\alpha, \beta, \gamma)$ and whose radius is equal to r . For, if $M(x, y, z)$ is an arbitrary point, then $\sqrt{(x - \alpha)^2 + (y - \beta)^2 + (z - \gamma)^2} = CM$. Hence, it is evident that equation (1) is satisfied by the coordinates of those points, and those only, whose distance from the point C is equal to r . Consequently, the locus of the points whose coordinates satisfy this equation is a sphere with centre $C(\alpha, \beta, \gamma)$ and radius r .

§ 60. The Equations of a Curve. The Problem of the Intersection of Three Surfaces

191. In solid analytic geometry every curve is regarded as the intersection of two surfaces and, accordingly, is *represented by two equations*.

For, if $F(x, y, z) = 0$ and $\Phi(x, y, z) = 0$ are the equations of two surfaces intersecting in a curve L , then the curve L is the locus of the points common to the two surfaces, that is, of the

points whose coordinates satisfy both equations $F(x, y, z) = 0$ and $\Phi(x, y, z) = 0$ simultaneously.

Thus, the two equations

$$\left. \begin{aligned} F(x, y, z) &= 0, \\ \Phi(x, y, z) &= 0, \end{aligned} \right\}$$

taken simultaneously, represent the curve L .

For example, the two equations

$$\left. \begin{aligned} (x-1)^2 + (y-2)^2 + (z-3)^2 &= 14, \\ x^2 + y^2 + z^2 &= 1, \end{aligned} \right\}$$

considered simultaneously, represent a circle (as the intersection of two spheres).

192. If $F(x, y, z) = 0$, $\Phi(x, y, z) = 0$, $\Psi(x, y, z) = 0$ are the equations of three surfaces, then each common solution of the system

$$\left. \begin{aligned} F(x, y, z) &= 0, \\ \Phi(x, y, z) &= 0, \\ \Psi(x, y, z) &= 0 \end{aligned} \right\}$$

gives the coordinates of a point common to the three surfaces. Consequently, *the geometric problem of finding the points of intersection of three surfaces is equivalent to the algebraic problem of solving simultaneously a system of three equations in three unknowns.*

Example. Find the points of intersection of three surfaces, given that the first surface is a sphere with centre at $(-1, -1, 0)$ and radius 5, the second surface a sphere with centre at $(1, 1, 3)$ and radius 4, and the third surface a plane parallel to the plane Oxy and situated in the upper half-space, 3 units above the plane Oxy .

Solution. The problem amounts to finding the simultaneous solutions of the three equations

$$\left. \begin{aligned} (x+1)^2 + (y+1)^2 + z^2 &= 25, \\ (x-1)^2 + (y-1)^2 + (z-3)^2 &= 16, \\ z &= 3. \end{aligned} \right\}$$

Substituting $z = 3$ in the first two equations and removing the parentheses, we get

$$\begin{aligned} x^2 + y^2 + 2x + 2y &= 14, \\ x^2 + y^2 - 2x - 2y &= 14. \end{aligned}$$

Hence $x + y = 0$, $x^2 + y^2 = 14$, so that $x = \pm\sqrt{7}$, $y = -x$. Thus, we obtain two points: $(\sqrt{7}, -\sqrt{7}, 3)$ and $(-\sqrt{7}, \sqrt{7}, 3)$.

§ 61. The Equation of a Cylindrical Surface with Elements Parallel to a Coordinate Axis

193. In this article we shall specially consider an equation of the form $F(x, y) = 0$. The distinctive feature of this equation is that its left-hand member lacks the variable z . This means that the equation relates only the first two coordinates, leaving the third coordinate free to assume any values.

We propose to show that *an equation of this form represents a cylindrical surface whose elements are parallel to the axis Oz .*

Let S denote the surface represented by an equation of the form $F(x, y) = 0$. Let $M_0(x_0, y_0, z_0)$ be an arbitrary point of the surface S . Since the point M_0 lies on S , it follows that the numbers x_0, y_0, z_0 satisfy the equation $F(x, y) = 0$; but then the numbers x_0, y_0, z , where z is any number whatsoever, also satisfy the equation, since $F(x, y)$ does not depend on z . Hence, for any z , the point $M(x_0, y_0, z)$ lies on the surface S (Fig. 102), which means that the straight line drawn through M_0 parallel to the axis Oz lies entirely on the surface S . Thus, the surface S is made up of straight lines parallel to the axis Oz , i. e., S is a cylindrical surface with elements parallel to the axis Oz , as was to be shown.

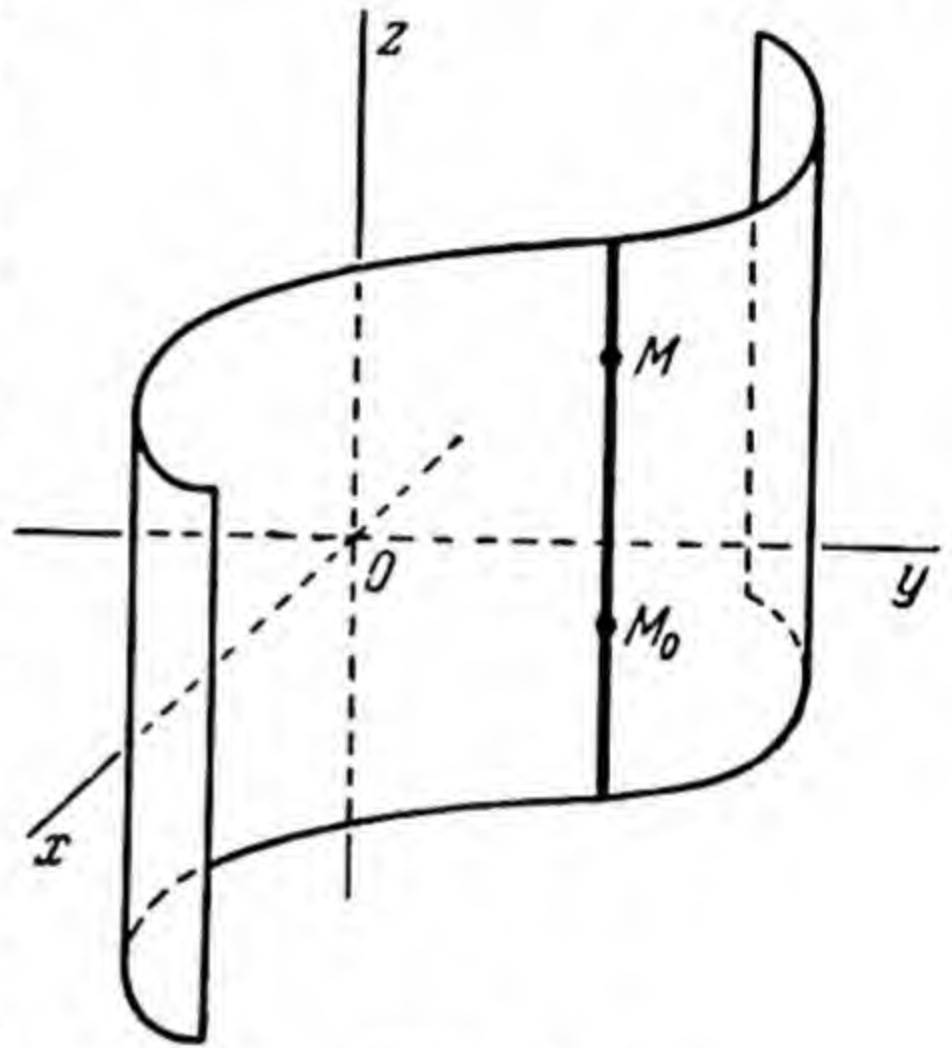


Fig. 102.

It should be observed that, in the plane Oxy (in the plane coordinate system determined by the axes Ox and Oy) the equation $F(x, y) = 0$ represents a *curve*, namely, the directing curve of the cylinder under consideration. In the *space* coordinate system, however, the same curve must be represented by two equations:

$$\left. \begin{aligned} F(x, y) &= 0, \\ z &= 0. \end{aligned} \right\}$$

Example. The equation $x^2 + y^2 = r^2$, referred to a space coordinate system, represents a circular cylinder; its directing curve (a circle), lying in the plane Oxy , is represented by the two equations

$$\left. \begin{aligned} x^2 + y^2 &= r^2, \\ z &= 0. \end{aligned} \right\}$$

194. By analogy with the above, it is easily seen that the equation $F(x, z) = 0$ represents (in space) a cylindrical surface with elements parallel to the axis Oy ; the equation $F(y, z) = 0$ represents a cylindrical surface with elements parallel to the axis Ox .

195. Consider a curve L (in space) represented by the equations

$$\left. \begin{aligned} F(x, y, z) &= 0, \\ \Phi(x, y, z) &= 0. \end{aligned} \right\} \quad (1)$$

Let

$$\Psi(x, y) = 0 \quad (2)$$

be the equation obtained from the system (1) by eliminating the variable z . This means that:

(1) equation (2) is a *consequence* of the system (1); i. e., each time the two equations of the system (1) are simultaneously satisfied by three numbers x, y, z , the first two of these numbers satisfy equation (2);

(2) if two numbers x, y satisfy equation (2), a third number z can be found such that the three numbers x, y, z will satisfy both equations of the system (1).

By Art. 193, equation (2) represents a cylindrical surface with elements parallel to the axis Oz . Further, in accordance with the first of the properties of equation (2) just stated, every point of the curve L lies on this cylindrical surface, which means that the surface passes through the curve L . Finally, according to the second property, each element of this surface passes through some point of the curve L . From all this, we conclude that the surface represented by the equation $\Psi(x, y) = 0$ is made up of straight lines which project the points of the curve L on the plane Oxy ; the surface is therefore called the cylindrical surface projecting the curve L on the plane Oxy (or simply the projecting cylinder).

The projection of the curve L on the plane Oxy is represented by two equations:

$$\left. \begin{aligned} \Psi(x, y) &= 0, \\ z &= 0. \end{aligned} \right\}$$

In a similar way, by eliminating the variable x or the variable y from the system (1), we can obtain the projection of the curve L on the plane Oyz or on the plane Oxz .

Example. The intersection of two spheres determines the circle

$$\left. \begin{aligned} x^2 + y^2 + z^2 &= 1, \\ x^2 + (y-1)^2 + (z-1)^2 &= 1. \end{aligned} \right\} \quad (3)$$

Find the projection of this circle on the plane Oxy .

Solution. We must find the equation of the cylinder projecting the given circle on the plane Oxy . This is achieved by eliminating z between the equations (3). Subtracting the second equation from the first, we get

$$y + z = 1; \quad (4)$$

hence $z = 1 - y$. Substituting the expression $1 - y$ for z in either of the given equations, we find

$$x^2 + 2y^2 - 2y = 0. \quad (5)$$

This is the desired result of the elimination of z from the system (3). For, equation (5) is a consequence of equations (3). Furthermore, if x and y satisfy equation (5), then the first of equations (3) gives

$$z = \pm \sqrt{1 - x^2 - y^2} = \pm \sqrt{1 + 2y^2 - 2y - y^2} = \pm (1 - y);$$

from the second equation of the system (3), we have

$$z - 1 = \pm \sqrt{1 - x^2 - (y-1)^2} = \pm \sqrt{1 + 2y^2 - 2y - y^2 + 2y - 1} = \pm y.$$

Thus, if two numbers x, y satisfy (5), then a third number z (namely, $z = 1 - y$) can be found such that the three numbers x, y, z will satisfy both equations of the system (3).

We see that the two conditions (see the beginning of this article) which must be satisfied by the result of the elimination of z from (3), are fulfilled for equation (5). By the foregoing, equation (5) represents the cylinder projecting the given circle on the plane Oxy . The projection itself is represented by the two equations

$$\left. \begin{aligned} x^2 + 2y^2 - 2y &= 0, \\ z &= 0. \end{aligned} \right\}$$

Since the first equation is reducible to the form $\frac{x^2}{\frac{1}{2}} + \frac{\left(y - \frac{1}{2}\right)^2}{\frac{1}{4}} = 1$, this

projection is an ellipse with semi-axes $a = \frac{1}{\sqrt{2}}, b = \frac{1}{2}$.

§ 62. Algebraic Surfaces

196. Solid analytic geometry has as its main subject of study the surfaces represented, in rectangular cartesian coordinates, by algebraic equations. These are equations of the following forms:

$$Ax + By + Cz + D = 0; \quad (1)$$

$$\begin{aligned} Ax^2 + By^2 + Cz^2 + 2Dxy + 2Exz + 2Fyz + 2Gx + \\ + 2Hy + 2Kz + L = 0; \quad (2) \end{aligned}$$

.....

A, B, C, D, E , etc., denote here fixed numbers and are called the coefficients of these equations.

Equation (1) is termed *the general equation of the first degree* (its coefficients may have any values whatsoever, provided only that the equation does contain first-degree terms; that is, A, B, C cannot all be zero at the same time); equation (2) is termed *the general equation of the second degree* (its coefficients may have any values whatsoever, provided only that the equation does contain second-degree terms, which means that the six coefficients A, B, C, D, E, F cannot all be zero at the same time). Equations of the third, fourth, etc., degrees have analogous forms.

A surface represented, in a rectangular cartesian system of coordinates, by an algebraic equation of degree n is called an algebraic surface of the n th order.

It can be proved that a surface, represented by an algebraic equation of degree n in *any* rectangular cartesian system of coordinates, will be represented in *any other* rectangular cartesian system by another algebraic equation of *the same degree n* . The proof, which is similar to that of Theorem 8 (Art. 49), is based on the formulas for transformation of rectangular cartesian coordinates in space.

197. The general theory of algebraic curves forms the subject of special treatises on analytic geometry. In this book, we are only concerned with surfaces of the first and the second order.

**THE PLANE AS THE SURFACE
OF THE FIRST ORDER.
THE EQUATIONS OF A STRAIGHT LINE**

**§ 63. The Plane as the Surface
of the First Order**

The next few sections are devoted to the establishment of the fact that surfaces of the first order are planes and only planes, and to the consideration of various forms of the equation of a plane.

198. Theorem 24. *Every plane is represented by an equation of the first degree in cartesian coordinates.*

Proof. Assuming that a rectangular cartesian system of coordinates has been attached to space, let us consider an arbitrary plane α and prove that this plane is represented by an equation of the first degree. Take a point $M_0(x_0, y_0, z_0)$ in the plane α ; in addition, choose any vector (except the zero vector!) *perpendicular* to the plane α . Denote the chosen vector by the letter \mathbf{n} , and its projections on the coordinate axes by A, B, C .

Let $M(x, y, z)$ be an arbitrary point. It lies in the plane α if, and only if, the vector $\overline{M_0M}$ is perpendicular to the vector \mathbf{n} . In other words, a point M lying in the plane α is characterised by the condition

$$\overline{M_0M} \perp \mathbf{n}.$$

Expressing this condition in terms of the coordinates x, y, z , we shall obtain the equation of the plane α . For this purpose, we write the coordinates of the vectors $\overline{M_0M}$ and \mathbf{n} .

$$\begin{aligned}\overline{M_0M} &= [x - x_0, y - y_0, z - z_0], \\ \mathbf{n} &= [A, B, C].\end{aligned}$$

By Art. 165, two perpendicular vectors have their scalar product zero; that is, the sum of the products of the corresponding coordinates of two perpendicular vectors is equal to zero. Hence, $\overline{M_0M} \perp \mathbf{n}$ if, and only if,

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \quad (1)$$

This is the desired equation of the plane α , since it is satisfied by the coordinates x, y, z of a point M if, and only if, M lies in the plane α (that is, if $\overline{M_0M} \perp \mathbf{n}$).

Removing the parentheses, we may write equation (1) in the form

$$Ax + By + Cz + (-Ax_0 - By_0 - Cz_0) = 0.$$

Denoting now the number $-Ax_0 - By_0 - Cz_0$ by D , we obtain

$$Ax + By + Cz + D = 0. \quad (2)$$

We see that a plane α is actually represented by an equation of the *first* degree. The proof is thus complete.

199. Every (non-zero) vector perpendicular to a plane is called its normal vector. Employing this term, we can say that the equation

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

is the equation of the plane passing through the point $M_0(x_0, y_0, z_0)$ and having $\mathbf{n} = \{A, B, C\}$ as its normal vector.

An equation of the form

$$Ax + By + Cz + D = 0$$

is called the *general equation* of a plane.

200. Theorem 25. Every equation of the first degree in cartesian coordinates represents a plane.

Proof. Assuming that a rectangular cartesian system of coordinates has been chosen, let us consider an arbitrary equation of the first degree:

$$Ax + By + Cz + D = 0. \quad (2)$$

By an "arbitrary" equation we mean that its coefficients A, B, C, D may be any numbers whatsoever (but, of course, excluding the case where the three coefficients A, B, C are all zero simultaneously). We must prove that equation (2) is the equation of a plane.

Let x_0, y_0, z_0 constitute a solution of equation (2), that is, let x_0, y_0, z_0 be a triad of numbers satisfying that equation^{*)}. Substituting the numbers x_0, y_0, z_0 for the current coordinates in the left-hand member of (2), we obtain the arithmetical identity

$$Ax_0 + By_0 + Cz_0 + D = 0. \quad (3)$$

^{*)} Equation (2), like every equation of the first degree in three unknowns, has infinitely many solutions. To obtain one of these, two unknowns are assigned arbitrary values, and then the third unknown is found from the equation.

Subtracting identity (3) from equation (2), we get the equation

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0, \quad (1)$$

which, according to the preceding article, is the equation of the plane passing through the point $M_0(x_0, y_0, z_0)$ and having $\mathbf{n} = \{A, B, C\}$ as its normal vector. Now, equation (2) is equivalent to equation (1), since equation (1) can be obtained from (2) by the term-by-term subtraction of identity (3), and since equation (2) can, in its turn, be obtained from (1) by the term-by-term addition of identity (3). Consequently, equation (2) is an equation of the same plane.

We have shown that an arbitrary equation of the *first degree* represents a *plane*; the theorem is thus proved.

201. As we know, surfaces represented (in cartesian coordinates) by equations of the first degree are termed surfaces of the first order. Using this term, we may express the above results as follows:

Every plane is a surface of the first order; every surface of the first order is a plane.

Example. Write the equation of the plane passing through the point $M_0(1, 1, 1)$ perpendicular to the vector $\mathbf{n} = \{2, 2, 3\}$.

Solution. By Art. 199, the required equation is

$$2(x - 1) + 2(y - 1) + 3(z - 1) = 0,$$

or

$$2x + 2y + 3z - 7 = 0.$$

202. To conclude the section, we shall prove the following proposition: *If the two equations $A_1x + B_1y + C_1z + D_1 = 0$ and $A_2x + B_2y + C_2z + D_2 = 0$ represent the same plane, their coefficients are proportional.*

For, in this case, the vectors $\mathbf{n}_1 = \{A_1, B_1, C_1\}$ and $\mathbf{n}_2 = \{A_2, B_2, C_2\}$ are perpendicular to the same plane and, hence, are collinear. But then, by Art. 154, the numbers A_2, B_2, C_2 are proportional to the numbers A_1, B_1, C_1 ; denoting the factor of proportionality by μ , we have $A_2 = A_1\mu$, $B_2 = B_1\mu$, $C_2 = C_1\mu$. Let $M_0(x_0, y_0, z_0)$ be any point in the plane; the coordinates of M_0 must satisfy each of the given equations, and so $A_1x_0 + B_1y_0 + C_1z_0 + D_1 = 0$ and $A_2x_0 + B_2y_0 + C_2z_0 + D_2 = 0$. Multiplying the first of these relations by μ and then subtracting it from the second relation, we obtain $D_2 - D_1\mu = 0$. Consequently, $D_2 = D_1\mu$ and

$$\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1} = \frac{D_2}{D_1} = \mu.$$

This proves our proposition.

§ 64. Incomplete Equations of Planes.

The Intercept Form of the Equation of a Plane

203. We know that *every* first-degree equation

$$Ax + By + Cz + D = 0$$

(in cartesian coordinates) represents a plane. Let us now consider some special cases of the equation of the first degree, namely, the cases where some of the coefficients A, B, C, D are zero.

(1) $D = 0$; the equation has the form $Ax + By + Cz = 0$ and represents a plane passing through the origin.

For, the numbers $x = 0, y = 0, z = 0$ satisfy the equation $Ax + By + Cz = 0$. Consequently, our plane contains the origin of coordinates.

(2) $C = 0$; the equation assumes the form $Ax + By + D = 0$ and represents a plane parallel to the axis Oz (or passing through that axis).

For, in this case, the normal vector $\mathbf{n} = \{A, B, C\}$ has its projection on the axis Oz equal to zero ($C = 0$); hence, the vector \mathbf{n} is perpendicular to the axis Oz , and the plane itself is parallel to Oz (or passes through Oz).

(3) $B = 0$ and $C = 0$; the equation has the form $Ax + D = 0$ and represents a plane parallel to the coordinate plane Oyz (or coincident with Oyz).

For, in this case, the normal vector $\mathbf{n} = \{A, B, C\}$ has its projections on the axes Oy and Oz equal to zero ($B = 0$ and $C = 0$); hence, the vector \mathbf{n} is perpendicular to the axes Oy and Oz , and the plane itself is parallel to these axes (or passes through each of them). But this means that the plane represented by the equation $Ax + D = 0$ is parallel to the plane Oyz or coincides with it. The same can also be verified in a different way, thus: rewrite the equation $Ax + D = 0$ as $x = -\frac{D}{A}$ and let

$$-\frac{D}{A} = a; \text{ this gives}$$

$$x = a.$$

According to this equation, all points of our plane have the same abscissa ($x = a$) and hence are all situated at the same distance from the plane Oyz (in "front" of it if $a > 0$, or in "back" of it if $a < 0$); consequently, a plane represented by such an equation is parallel to the plane Oyz . From this, it is also clear that a is the intercept of our plane on the axis Ox . In particular, when $D = 0$, then $a = 0$; in this case, the plane under

consideration coincides with the plane Oyz . Thus, the equation $x = 0$ represents the plane Oyz .

204. By analogy with the above, it is easily established that:

1. An equation of the form $Ax + Cz + D = 0$ represents a plane parallel to the axis Oy (or passing through it); an equation of the form $By + Cz + D = 0$ represents a plane parallel to the axis Ox (or passing through it).

2. An equation of the form $By + D = 0$ represents a plane parallel to the plane Oxz (or coincident with it); an equation of the form $Cz + D = 0$ represents a plane parallel to the plane Oxy (or coincident with it). The last two equations can be written, respectively, as $y = b$ and $z = c$, where b is the intercept of the first plane on the axis Oy , and c is the intercept of the second plane on the axis Oz . In particular, the equation $y = 0$ represents the plane Oxz , and the equation $z = 0$ represents the plane Oxy .

205. Let us consider the equation of a plane,

$$Ax + By + Cz + D = 0,$$

in which all the coefficients A, B, C, D are different from zero. This equation can be reduced to a special form, which is found convenient when dealing with some problems of analytic geometry.

Transposing the constant term D to the right-hand side of the equation, we obtain

$$Ax + By + Cz = -D.$$

Dividing both sides of the equation by $-D$, we then obtain

$$\frac{Ax}{-D} + \frac{By}{-D} + \frac{Cz}{-D} = 1,$$

or

$$-\frac{x}{\frac{D}{A}} + -\frac{y}{\frac{D}{B}} + -\frac{z}{\frac{D}{C}} = 1.$$

Introducing the notation

$$a = -\frac{D}{A}, \quad b = -\frac{D}{B}, \quad c = -\frac{D}{C},$$

we have

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1. \quad (1)$$

This is the special form of the equation of a plane that we wished to obtain. The numbers a, b, c have here a very simple geometric meaning; namely, a, b and c are the intercepts of the plane on the respective coordinate axes. To verify this, let us

find the points in which our plane is pierced by the coordinate axes. The point of intersection of the plane and the axis Ox is determined from the equation of the plane, $\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$, under the additional condition $y=0, z=0$; hence $x=a$, and so the intercept of the plane on the axis Ox is actually equal to a . In like manner, the intercepts of the plane on the axes Oy and Oz are shown to be equal to b and c , respectively.

An equation of the form (1) is referred to as *the intercept equation of a plane*.

Example. Find the equation of the plane whose intercepts on the coordinate axes are $a = 2, b = -3, c = 4$.

Solution. By the foregoing, the desired equation can be written at once: $\frac{x}{2} - \frac{y}{3} + \frac{z}{4} = 1$, or $6x - 4y + 3z - 12 = 0$.

§ 65. The Normal Equation of a Plane. The Distance of a Point from a Plane

206. We shall now consider still another special form, known as the *normal* form, of the equation of a plane.

Let there be given any plane π . Through the origin, draw the straight line n (called the *normal*) perpendicular to the plane π , and denote by P the point in which the normal pierces the plane π (Fig. 103). We shall regard the direction from the point O to the point P on the normal as its positive direction (if P coincides with O , that is, if the given plane passes through the origin, the positive direction of the normal may be chosen at will). Let α, β, γ denote the angles which the directed normal makes with the coordinate axes, and let p be the length of the segment OP .

We proceed to derive the equation of the given plane π , assuming that the numbers $\cos \alpha, \cos \beta, \cos \gamma$ and p are known. For this purpose, take an arbitrary point M in the plane π and designate the coordinates of M as x, y, z . The projection of the vector \overline{OM} on the normal is clearly equal to OP and, since the positive direction of the normal agrees with that of the segment \overline{OP} , the value of this segment is represented by a positive number, namely, the number p ; thus,

$$\text{proj}_n \overline{OM} = p. \quad (1)$$

Now observe that $\overline{OM} = \{x, y, z\}$. Hence, by the third corollary to Theorem 20 (Art. 165),

$$\text{proj}_n \overline{OM} = x \cos \alpha + y \cos \beta + z \cos \gamma. \quad (2)$$

From relations (1) and (2), it follows that $x \cos \alpha + y \cos \beta + z \cos \gamma = p$, or

$$x \cos \alpha + y \cos \beta + z \cos \gamma - p = 0. \quad (3)$$

This is the desired equation of the given plane (as we can see, it is satisfied by the coordinates x, y, z of every point M lying in the plane; on the other hand, if a point M does not lie in the plane,

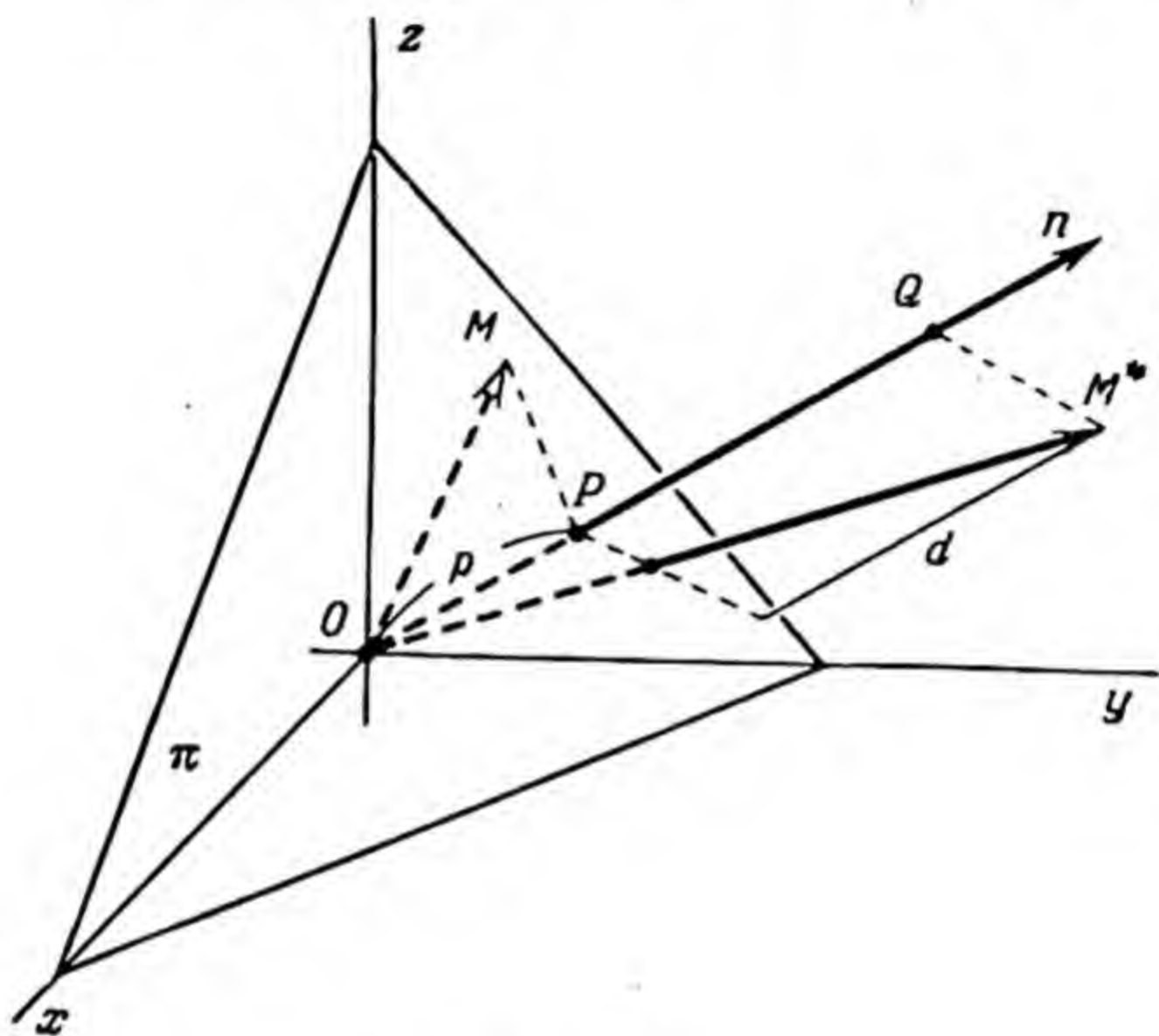


Fig. 103.

the coordinates of M do not satisfy equation (3), because then $\text{proj}_n \overline{OM} \neq p$).

The equation of a plane written in the form (3) is called the *normal equation of a plane*; $\cos \alpha, \cos \beta, \cos \gamma$ are here the *direction cosines of the normal*, and p is the *distance of the plane from the origin*.

207. Let there be given an arbitrary plane. Draw its normal n and assign a positive direction to the normal as described in the preceding article. Further, let M^* be any point in space, and let d denote the distance of M^* from the given plane (see Fig. 103).

We shall agree to define the *departure* of the point M^* from the given plane as the number $+d$ if M^* lies on that side of the plane towards which the directed normal points, and as the number $-d$ if M^* lies on the other side of the plane. We shall denote

the departure of a point from a plane by the letter δ ; thus, $\delta = \pm d$, and it will be helpful to note that $\delta = +d$ when the point M^* and the origin are *on opposite sides* of the plane, and $\delta = -d$ when M^* and the origin are *on the same side* of the plane. (For points lying in the plane, $\delta = 0$.)

Theorem 26. *For a point M^* having coordinates (x^*, y^*, z^*) and a plane represented by the normal equation*

$$x \cos \alpha + y \cos \beta + z \cos \gamma - p = 0,$$

the departure of the point M^ from the plane is given by the formula*

$$\delta = x^* \cos \alpha + y^* \cos \beta + z^* \cos \gamma - p. \quad (4)$$

Proof. Project the point M^* on the normal; let Q be the projection (see Fig. 103); then

$$\delta = PQ = OQ - OP,$$

where PQ , OQ and OP are the values of the directed segments \overline{PQ} , \overline{OQ} and \overline{OP} of the normal. But $OQ = \text{proj}_n \overline{OM^*}$, $OP = p$; hence

$$\delta = \text{proj}_n \overline{OM^*} - p. \quad (5)$$

By the third corollary to Theorem 20 (Art. 165),

$$\text{proj}_n \overline{OM^*} = x^* \cos \alpha + y^* \cos \beta + z^* \cos \gamma. \quad (6)$$

From (5) and (6), we obtain

$$\delta = x^* \cos \alpha + y^* \cos \beta + z^* \cos \gamma - p.$$

The theorem is thus proved.

Note now that $x^* \cos \alpha + y^* \cos \beta + z^* \cos \gamma - p$ is nothing more than the left-hand member of the normal equation of the given plane, with the current coordinates replaced by the coordinates of the point M^* . Hence we have the following rule:

To find the departure of a point M^ from a plane, the coordinates of the point M^* must be substituted for the current coordinates in the left-hand member of the normal equation of the plane. The resulting number will be the required departure.*

Note. To find the *distance* of a point from a plane, we have merely to compute the *departure* by the rule just given, and to take its *absolute* value.

208. We shall now show how to reduce the *general* equation of a plane to the normal form. Let

$$Ax + By + Cz + D = 0 \quad (7)$$

be the general equation of a plane, and let

$$x \cos \alpha + y \cos \beta + z \cos \gamma - p = 0 \quad (3)$$

be its normal equation.

Since equations (7) and (3) represent the same plane, the coefficients of these equations are, by Art. 202, proportional. This means that, on multiplying equation (7) throughout by a certain factor μ , we shall obtain the equation

$$\mu Ax + \mu By + \mu Cz + \mu D = 0,$$

which will be identical with equation (3); that is, we shall have

$$\mu A = \cos \alpha, \quad \mu B = \cos \beta, \quad \mu C = \cos \gamma, \quad \mu D = -p. \quad (8)$$

To find the factor μ , square and add the first three of these relations; this gives

$$\mu^2 (A^2 + B^2 + C^2) = \cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma.$$

But, according to Art. 140, the right member of this last relation is equal to unity. Hence

$$\mu = \frac{\pm 1}{\sqrt{A^2 + B^2 + C^2}}. \quad (9)$$

The number μ , multiplication by which reduces the general equation of a plane to the normal form, is called the *normalising factor* of that equation. Formula (9) determines the *normalising factor* incompletely, since its sign remains unspecified. To determine the sign of the normalising factor, let us use the fourth of relations (8). According to this relation, $\mu D = -p$, which means that μD is a negative number.

Hence, *the normalising factor is opposite in sign to the constant term of the equation to be normalised.*

Note. If $D = 0$, the sign of the normalising factor may be chosen at pleasure.

Example. Given the plane $3x - 4y + 12z + 14 = 0$ and the point $M(4, 3, 1)$. Find the departure of M from the plane.

Solution. To apply the rule stated in Art. 207, we must first reduce the given equation to its normal form. For this purpose, we find the normalising factor

$$\mu = \frac{-1}{\sqrt{3^2 + 4^2 + 12^2}} = -\frac{1}{13}.$$

Multiplying our equation by μ , we get the desired normal equation of the plane:

$$-\frac{1}{13} (3x - 4y + 12z + 14) = 0.$$

Substituting the coordinates of M in the left member of this equation, we have

$$\delta = -\frac{1}{13} (3 \cdot 4 - 4 \cdot 3 + 12 \cdot 1 + 14) = -2.$$

Thus, the point M has a negative departure from the given plane and is at the distance $d = 2$ from that plane.

§ 66. The Equations of a Straight Line

209. We have already pointed out in Art. 191 that, in solid analytic geometry, every curve is regarded as the intersection of two surfaces and is represented by two equations. In particular, every *straight line* will be regarded as the intersection of two *planes* and will, accordingly, be represented by two equations of the *first* degree (in cartesian coordinates, of course).

Let a rectangular cartesian system of coordinates be attached to space. Consider an arbitrary straight line a (Fig. 104). Let π_1

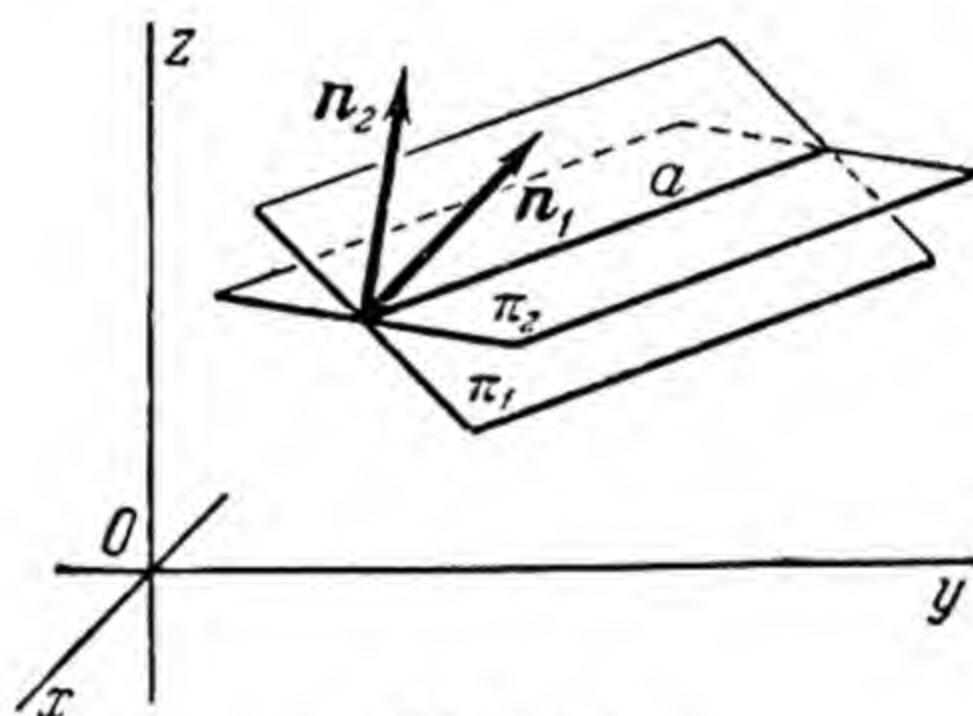


Fig. 104.

and π_2 denote any two distinct planes intersecting in the line a , and suppose that the equations of these planes are known; we shall write them as

$$A_1x + B_1y + C_1z + D_1 = 0 \quad \text{and} \quad A_2x + B_2y + C_2z + D_2 = 0.$$

Since the line a is the intersection of the planes π_1 and π_2 , it is represented by their two equations, taken simultaneously:

$$\left. \begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0. \end{aligned} \right\} \quad (1)$$

210. Suppose that we are given beforehand two equations of the first degree; let them be of the form (1). Will they, when considered simultaneously, always represent a straight line? Clearly, they will represent a straight line if, and only if, the correspond-

ing planes are non-parallel and non-coincident, that is, if the normal vectors $\mathbf{n}_1 = \{A_1, B_1, C_1\}$ and $\mathbf{n}_2 = \{A_2, B_2, C_2\}$ of these planes are non-collinear. Recalling that the collinearity condition for two vectors is that their coordinates should be proportional (see Art. 154), we conclude:

Two equations of the form (1), taken simultaneously, represent a straight line if, and only if, the coefficients A_1, B_1, C_1 of one of the equations are not proportional to the coefficients A_2, B_2, C_2 of the other.

211. An infinite number of distinct planes pass through every straight line; obviously, there exist infinitely many possibilities of selecting two planes from that number. It follows that every straight line can be represented by two equations chosen in infinitely many different ways. We shall now show a very simple method which enables us, *by combining the known equations of two planes passing through the given straight line, to obtain from them any desired number of new equations, each of which will also represent a plane passing through the given line.*

Let there be given a straight line a and the equations of two distinct planes π_1 and π_2 , passing through this line:

$$A_1x + B_1y + C_1z + D_1 = 0 \quad \text{and} \quad A_2x + B_2y + C_2z + D_2 = 0.$$

Let us take any two numbers α and β , which are not both simultaneously equal to zero, and form the relation

$$\alpha(A_1x + B_1y + C_1z + D_1) + \beta(A_2x + B_2y + C_2z + D_2) = 0, \quad (2)$$

or, written in a different form,

$$(\alpha A_1 + \beta A_2)x + (\alpha B_1 + \beta B_2)y + (\alpha C_1 + \beta C_2)z + (\alpha D_1 + \beta D_2) = 0. \quad (3)$$

It is easy to show that the three numbers $\alpha A_1 + \beta A_2$, $\alpha B_1 + \beta B_2$ and $\alpha C_1 + \beta C_2$ cannot all be zero simultaneously. For, if $\alpha A_1 + \beta A_2 = 0$, $\alpha B_1 + \beta B_2 = 0$, $\alpha C_1 + \beta C_2 = 0$, then

$$\frac{A_1}{A_2} = -\frac{\beta}{\alpha}, \quad \frac{B_1}{B_2} = -\frac{\beta}{\alpha}, \quad \frac{C_1}{C_2} = -\frac{\beta}{\alpha}.$$

Since the numbers α and β are not both zero, the ratio $\frac{\beta}{\alpha}$ cannot be indeterminate; from the above proportions, it therefore follows that A_1, B_1, C_1 are proportional to A_2, B_2, C_2 , i. e., that the normal vectors $\mathbf{n}_1 = \{A_1, B_1, C_1\}$ and $\mathbf{n}_2 = \{A_2, B_2, C_2\}$ of the given planes are collinear; but this is impossible, since the given planes are non-parallel and non-coincident.

Since the three numbers $\alpha A_1 + \beta A_2$, $\alpha B_1 + \beta B_2$, and $\alpha C_1 + \beta C_2$ cannot all vanish at the same time, relation (3) is an equation. It

is obviously an equation of the first degree and hence it represents a plane.

Furthermore, since equation (3) is a consequence of the equations $A_1x + B_1y + C_1z + D_1 = 0$ and $A_2x + B_2y + C_2z + D_2 = 0$, it follows that every triad of numbers (x, y, z) satisfying these two equations satisfies equation (3) as well. Hence, every point lying on the intersection of the planes π_1 and π_2 also lies in the plane represented by equation (3). In other words, equation (3)—or equation (2), which is equivalent to (3)—represents a plane through the line a .

Thus, if

$$A_1x + B_1y + C_1z + D_1 = 0 \quad \text{and} \quad A_2x + B_2y + C_2z + D_2 = 0$$

are the equations of two planes passing through a straight line, then the equation

$$\alpha(A_1x + B_1y + C_1z + D_1) + \beta(A_2x + B_2y + C_2z + D_2) = 0 \quad (2)$$

represents a plane passing through the same straight line.

We can use this proposition to simplify the equations of a straight line. For example, the equations of the straight line

$$\left. \begin{aligned} x + y + z + 3 &= 0, \\ x + y - z - 5 &= 0 \end{aligned} \right\}$$

can be replaced by simpler ones as follows: combine the given equations, setting first $\alpha = 1$, $\beta = 1$, and then $\alpha = 1$, $\beta = -1$; this will give the equations

$$\left. \begin{aligned} x + y - 1 &= 0, \\ z + 4 &= 0, \end{aligned} \right\}$$

representing the same straight line as the original equations.

212. Let a straight line a be represented by the equations

$$\left. \begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0, \end{aligned} \right\}$$

as the intersection of the two planes π_1 and π_2 . We know that the equation

$$\alpha(A_1x + B_1y + C_1z + D_1) + \beta(A_2x + B_2y + C_2z + D_2) = 0, \quad (2)$$

for all values of α , β (not both zero) represents a plane through the straight line a . We shall now prove that it is always possible to select the values of α , β so that equation (2) will represent any (previously assigned) plane passing through the line a .

Since each plane passing through the line a is determined by specifying (beside the line a) one of its points, it follows that, in order to prove the assertion just made, we have merely to show that the numbers α, β in (2) can always be chosen so as to make the plane represented by (2) pass through any preassigned point $M^*(x^*, y^*, z^*)$.

But this is evident; for, the plane represented by equation (2) will pass through a point M^* if the coordinates of M^* satisfy this equation, that is, if

$$\alpha(A_1x^* + B_1y^* + C_1z^* + D_1) + \beta(A_2x^* + B_2y^* + C_2z^* + D_2) = 0. \quad (4)$$

We assume that the point M^* does not lie on the line a (this being the only case we are concerned with). Then at least one of the numbers $A_1x^* + B_1y^* + C_1z^* + D_1, A_2x^* + B_2y^* + C_2z^* + D_2$ is different from zero, and hence (4) is an equation of the first degree in two unknowns, α and β . To find the unknowns α, β , one of them is assigned an arbitrary value, and then the value of the other is computed from the equation; for example, if $A_2x^* + B_2y^* + C_2z^* + D_2 \neq 0$, then α may be assigned any value (other than zero), and the corresponding value of β may then be determined from the relation

$$\beta = -\frac{A_1x^* + B_1y^* + C_1z^* + D_1}{A_2x^* + B_2y^* + C_2z^* + D_2} \alpha.$$

Thus, an equation of the form (2) can be made to represent a plane passing through any preassigned point in space, and hence to represent any plane passing through the given straight line a .

213. *The totality of planes passing through the same straight line is called a pencil of planes. An equation of the form (2) is called the equation of a pencil of planes since, by assigning appropriate values to α and β , it can be made to represent every plane of a pencil.*

If $\alpha \neq 0$, then, letting $\frac{\beta}{\alpha} = \lambda$, we obtain from (2)

$$A_1x + B_1y + C_1z + D_1 + \lambda(A_2x + B_2y + C_2z + D_2) = 0. \quad (5)$$

In practical solving of problems, this form of the equation of a pencil of lines is used more frequently than the form (2). However, it is important to note that, since the case $\alpha = 0$ is excluded when reducing (2) to (5), an equation of the form (5) cannot represent the plane $A_2x + B_2y + C_2z + D_2 = 0$; that is, an equation of the form (5) can be made, by varying the value of λ , to represent every plane of the pencil except one (the second of the two given planes).

**§ 67. The Direction Vector of a Straight Line.
The Canonical Equations of a Straight Line.
The Parametric Equations of a Straight Line**

214. We shall now introduce a special form of the equations of a straight line, which is conveniently used in solving some problems of analytic geometry. This special form of the equations of a straight line can be derived from its general equations by algebraic transformations; however, we prefer to establish it by a direct method, exhibiting the geometric aspect of the matter.

Let there be given a straight line. *Every non-zero vector lying on the given line or parallel to it is called the direction vector of that line.* The term "direction vector" is applied to such vectors because any one of them, once it has been specified, determines the direction of the line.

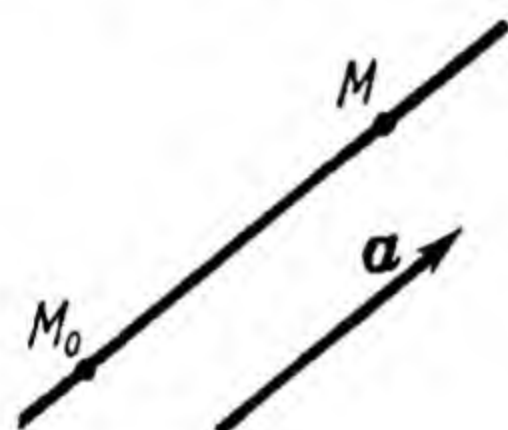


Fig. 105.

We shall denote the direction vector of an arbitrary straight line by the letter \mathbf{a} , and the coordinates of this vector by l, m, n :

$$\mathbf{a} = \{l, m, n\}.$$

We proceed now to derive the equations of the straight line passing through the given point $M_0(x_0, y_0, z_0)$ and having the given direction vector $\mathbf{a} = \{l, m, n\}$.

These equations are easily obtained as follows. Let $M(x, y, z)$ be an arbitrary ("variable") point of the straight line (Fig. 105). The vector

$$\overline{M_0M} = \{x - x_0, y - y_0, z - z_0\}$$

is collinear with the direction vector

$$\mathbf{a} = \{l, m, n\}.$$

Hence the coordinates of the vector $\overline{M_0M}$ are proportional to those of the vector \mathbf{a} :

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}. \quad (1)$$

We see that these relations are satisfied by the coordinates of every point $M(x, y, z)$ lying on the line under consideration; on the other hand, if a point $M(x, y, z)$ does not lie on the line, its coordinates do not satisfy relations (1), since in this case the vectors $\overline{M_0M}$ and \mathbf{a} are non-collinear and their coordinates are not in proportion. Thus, equations (1) are the equations of the

straight line passing through the point $M_0(x_0, y_0, z_0)$ in the direction of the vector $\mathbf{a} = \{l, m, n\}$.

We shall refer to equations having the special form just derived as *the canonical equations* of a straight line.

The coordinates l, m, n of any direction vector \mathbf{a} of a straight line are called *the direction parameters* of the line; the direction cosines of the vector \mathbf{a} are called *the direction cosines* of the line.

215. Let a straight line be represented by two *general* equations,

$$\left. \begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0. \end{aligned} \right\} \quad (2)$$

We shall now show how to derive the *canonical* equations of this line.

Let π_1 and π_2 denote the planes represented by the given equations, and let \mathbf{n}_1 and \mathbf{n}_2 be the normal vectors of these planes. To form the canonical equations of our straight line, it is necessary:

(1) to find a point $M_0(x_0, y_0, z_0)$ of this line by assigning an arbitrary value to one of the unknown coordinates x_0, y_0, z_0 and substituting this value for the corresponding variable in equations (2), whereupon the two other coordinates will be determined from equations (2) by solving them simultaneously;

(2) to find the direction vector $\mathbf{a} = \{l, m, n\}$. Since the given straight line is determined by the intersection of the planes π_1 and π_2 , it is perpendicular to each of the vectors \mathbf{n}_1 and \mathbf{n}_2 (see Fig. 104). Hence we can take as the vector \mathbf{a} any vector perpendicular to the vectors \mathbf{n}_1 and \mathbf{n}_2 , say, their vector product: $\mathbf{a} = [\mathbf{n}_1\mathbf{n}_2]$. Since the coordinates of the vectors \mathbf{n}_1 and \mathbf{n}_2 are known: $\mathbf{n}_1 = \{A_1, B_1, C_1\}$, $\mathbf{n}_2 = \{A_2, B_2, C_2\}$, the coordinates of the vector $\mathbf{a} = \{l, m, n\}$ can be computed by simply using Theorem 21 (Art. 173).

Example. Find the canonical equations of the straight line

$$\left. \begin{aligned} 3x + 2y + 4z - 11 &= 0, \\ 2x + y - 3z - 1 &= 0. \end{aligned} \right\}$$

Solution. Setting, for example, $x_0 = 1$, we find from the given system: $y_0 = 2$, $z_0 = 1$; thus, we have already determined a point $M_0(1, 2, 1)$ of the line. Let us now find its direction vector. We have: $\mathbf{n}_1 = \{3, 2, 4\}$, $\mathbf{n}_2 = \{2, 1, -3\}$; hence $\mathbf{a} = [\mathbf{n}_1\mathbf{n}_2] = \{-10, 17, -1\}$, that is, $l = -10$, $m = 17$, $n = -1$. The canonical equations of the line are obtained by substituting the values of x_0, y_0, z_0 and l, m, n so found in equations (1):

$$\frac{x-1}{-10} = \frac{y-2}{17} = \frac{z-1}{-1}.$$

216. Let there be given the canonical equations of a straight line. Denote by t each of the equal ratios forming part of these canonical equations; we then have

$$\frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n} = t.$$

Hence

$$\left. \begin{aligned} x &= x_0 + lt, \\ y &= y_0 + mt, \\ z &= z_0 + nt. \end{aligned} \right\} \quad (3)$$

These are the *parametric equations of the straight line passing through the point $M_0(x_0, y_0, z_0)$ in the direction of the vector $\mathbf{a} = \{l, m, n\}$* . In equation (3), t is regarded as an arbitrarily varying parameter, and x, y, z as functions of t ; as t varies, the quantities x, y, z vary in such a manner that the point $M(x, y, z)$ moves along the given straight line. The parametric equations of a straight line are conveniently used in cases where it is required to find the point of intersection of the straight line with a plane.

Example. Given the straight line

$$\frac{x - 2}{1} = \frac{y - 3}{1} = \frac{z - 4}{2}$$

and the plane $2x + y + z - 6 = 0$; find their point of intersection.

Solution. The problem amounts to determining x, y, z from the three given equations (we have two equations of the straight line and one equation of the plane). The necessary computations will be made simpler if we raise the number of unknowns (and the number of equations) to four by letting

$$\frac{x - 2}{1} = \frac{y - 3}{1} = \frac{z - 4}{2} = t; \text{ hence}$$

$$x = 2 + t, \quad y = 3 + t, \quad z = 4 + 2t.$$

Substituting these expressions in the left member of the equation of the given plane, we at once obtain a single equation in one unknown:

$$2(2 + t) + (3 + t) + (4 + 2t) - 6 = 0.$$

Solving this equation, we get $t = -1$, and hence the coordinates of the required point are $x = 1, y = 2, z = 2$.

217. Let us agree to regard t as the number of seconds that have passed from a preset instant of time ("the instant of starting the stop-watch"), and let us consider equations (3) as the *equations of motion* of the point $M(x, y, z)$ (see Art. 45). We now proceed to clarify the nature of this motion.

First of all, it is evident from the foregoing that the point M is in rectilinear motion along the straight line passing through the point M_0 in the direction of the vector $\mathbf{a} = \{l, m, n\}$.

Furthermore, it is easy to verify that the motion of the point M , as determined by equations (3), is *uniform* motion. For, by (3), we have

$$x - x_0 = lt, \quad y - y_0 = mt, \quad z - z_0 = nt;$$

these three relations are equivalent to the single vector equation

$$\overline{M_0M} = \mathbf{a}t.$$

It is hence apparent that the displacement $\overline{M_0M}$ experienced by the point M in the time of t seconds is equal to the vector \mathbf{a} elongated " t -fold". Thus, the displacement of the point M is proportional to time t , which means that the motion of the point M is uniform.

Finally, let us compute the *speed* of the point M . For this purpose, note that, in the course of the first second (from $t = 0$ till $t = 1$), the point M experiences the displacement $\overline{M_0M} = \mathbf{a}$. Consequently, the velocity of the point M is numerically equal to the modulus of the vector \mathbf{a} ; that is, the speed of M is $v = \sqrt{l^2 + m^2 + n^2}$. Thus, equations (3) determine the *uniform rectilinear motion* of the point $M(x, y, z)$ at the speed $v = \sqrt{l^2 + m^2 + n^2}$ in the direction of the vector $\mathbf{a} = \{l, m, n\}$; the point $M_0(x_0, y_0, z_0)$ is the initial position of the variable point $M(x, y, z)$, which means that the point M coincides with the point M_0 at $t = 0$.

Example. Find the equations of motion of the point $M(x, y, z)$ which, starting from $M_0(1, 1, 1)$, moves rectilinearly and uniformly in the direction of the vector $\mathbf{s} = \{2, 3, 6\}$, at the speed $v = 21$.

Solution. Comparing the modulus of the vector \mathbf{s} , which is equal to $\sqrt{2^2 + 3^2 + 6^2} = 7$, with the given speed $v = 21$, we see that we must take the vector \mathbf{s} stretched threefold as the vector \mathbf{a} ; that is, $\mathbf{a} = \{6, 9, 18\}$. The desired equations are

$$x = 1 + 6t, \quad y = 1 + 9t, \quad z = 1 + 18t.$$

§ 68. Some Additional Propositions and Examples

218. In analytic geometry it is often required to write the equations of a straight line two of whose points are given. We shall now find the general solution of this problem, letting

$$M_1(x_1, y_1, z_1) \quad \text{and} \quad M_2(x_2, y_2, z_2)$$

be two given arbitrary points of the line.

In order to solve the problem, it is sufficient to note that the vector $\mathbf{a} = \overline{M_1M_2}$ can be taken as the direction vector of the line in question; hence

$$l = x_2 - x_1, \quad m = y_2 - y_1, \quad n = z_2 - z_1.$$

Assigning to the point $M_1(x_1, y_1, z_1)$ the role played by the point M_0 in Art. 215, we obtain

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

These are the desired (canonical) *equations of the straight line passing through the two given points* $M_1(x_1, y_1, z_1)$ and $M_2(x_2, y_2, z_2)$.

219. Let us also find the general solution of the following problem: *To write the equation of the plane passing through the three distinct points* $M_1(x_1, y_1, z_1)$, $M_2(x_2, y_2, z_2)$ and $M_3(x_3, y_3, z_3)$.

Let x, y, z denote the coordinates of an arbitrary point M in space, and consider the three vectors $\overline{M_1M} = \{x - x_1, y - y_1, z - z_1\}$, $\overline{M_1M_2} = \{x_2 - x_1, y_2 - y_1, z_2 - z_1\}$ and $\overline{M_1M_3} = \{x_3 - x_1, y_3 - y_1, z_3 - z_1\}$. The point M lies in the plane $M_1M_2M_3$ if, and only if, the vectors $\overline{M_1M}$, $\overline{M_1M_2}$ and $\overline{M_1M_3}$ are coplanar; by Art. 185, the condition for coplanarity of these three vectors is that the determinant of the third order formed from their coordinates should be equal to zero. In the present case, this gives

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0.$$

This is the desired equation of the plane passing through the points M_1, M_2, M_3 , since it is satisfied by the coordinates x, y, z of a point M if, and only if, the point M lies in the plane.

220. The solution of a number of problems of analytic geometry requires the knowledge of *conditions for the parallelism and the perpendicularity* of two planes, of two straight lines, or of a straight line and a plane. We proceed to derive these conditions.

(1) Let

$$\begin{aligned} A_1x + B_1y + C_1z + D_1 &= 0, \\ A_2x + B_2y + C_2z + D_2 &= 0 \end{aligned}$$

be the equations of two given planes. These planes are parallel if, and only if, their normal vectors $\mathbf{n}_1 = \{A_1, B_1, C_1\}$, $\mathbf{n}_2 = \{A_2, B_2, C_2\}$ are collinear (Fig. 106; coincident planes are

considered here as a special case of parallel planes). Hence, by Art. 154, we obtain the following condition for the parallelism of two planes:

$$\frac{A_2}{A_1} = \frac{B_2}{B_1} = \frac{C_2}{C_1}.$$

The given planes are perpendicular if, and only if, their normal vectors are perpendicular (Fig. 107). Hence, by Art. 165, we have the following condition for the perpendicularity of two planes:

$$A_1A_2 + B_1B_2 + C_1C_2 = 0.$$

(2) Let

$$\frac{x - x_1}{l_1} = \frac{y - y_1}{m_1} = \frac{z - z_1}{n_1},$$

$$\frac{x - x_2}{l_2} = \frac{y - y_2}{m_2} = \frac{z - z_2}{n_2}$$

be the equations of two given straight lines. These lines are parallel if, and only if, their direction vectors $\mathbf{a}_1 = \{l_1, m_1, n_1\}$,

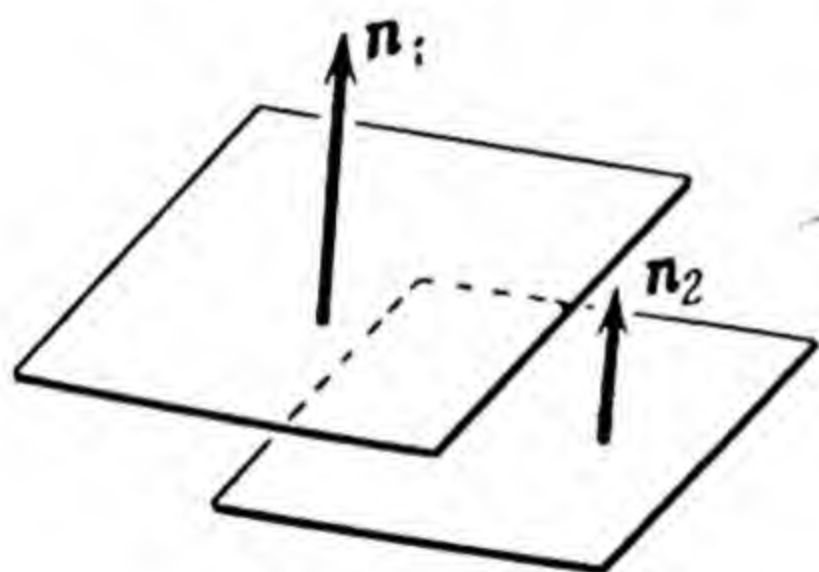


Fig. 106.

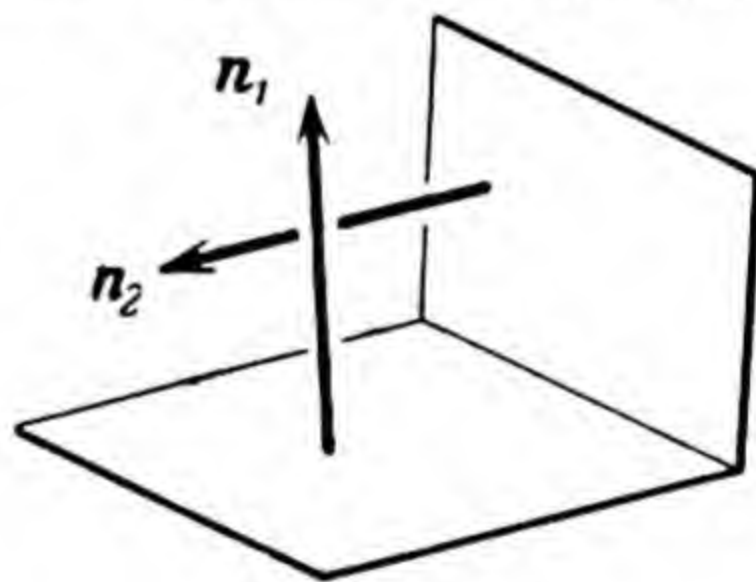


Fig. 107.

$\mathbf{a}_2 = \{l_2, m_2, n_2\}$ are collinear (Fig. 108; coincident lines are considered here as a special case of parallel lines). This gives the following condition for the parallelism of two straight lines:

$$\frac{l_2}{l_1} = \frac{m_2}{m_1} = \frac{n_2}{n_1}.$$

The given straight lines are perpendicular if, and only if, their direction vectors are perpendicular (Fig. 109; in space, perpendicular straight lines need not necessarily intersect). Hence we get the condition for the perpendicularity of two straight lines:

$$l_1l_2 + m_1m_2 + n_1n_2 = 0.$$

(3) Let there be given a straight line,

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{n},$$

and a plane,

$$Ax + By + Cz + D = 0.$$

The straight line is parallel to the plane if, and only if, the direction vector $\mathbf{a} = \{l, m, n\}$ of the line is perpendicular to the

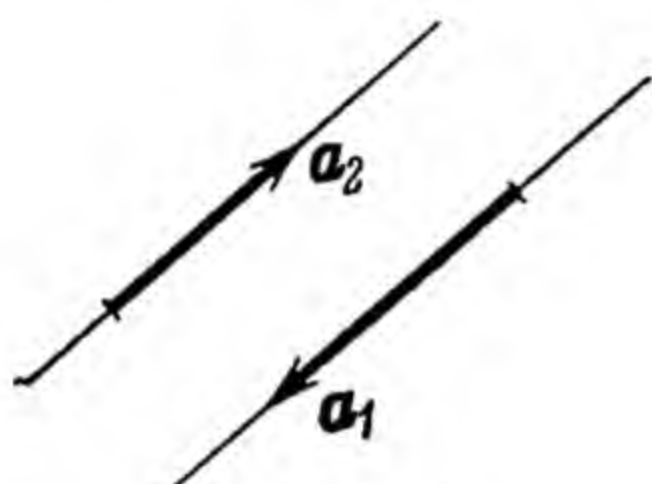


Fig. 108.

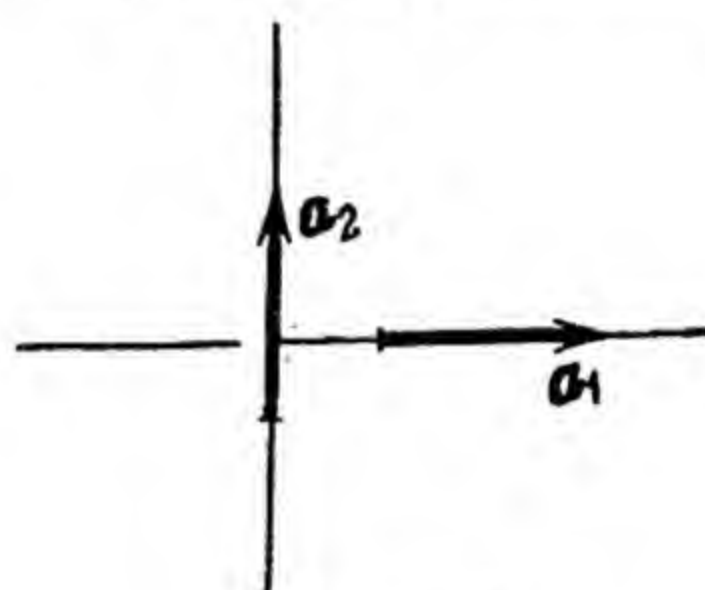


Fig. 109.

normal vector $\mathbf{n} = \{A, B, C\}$ of the plane (Fig. 110; the case when the straight line lies in the plane is considered here as a special

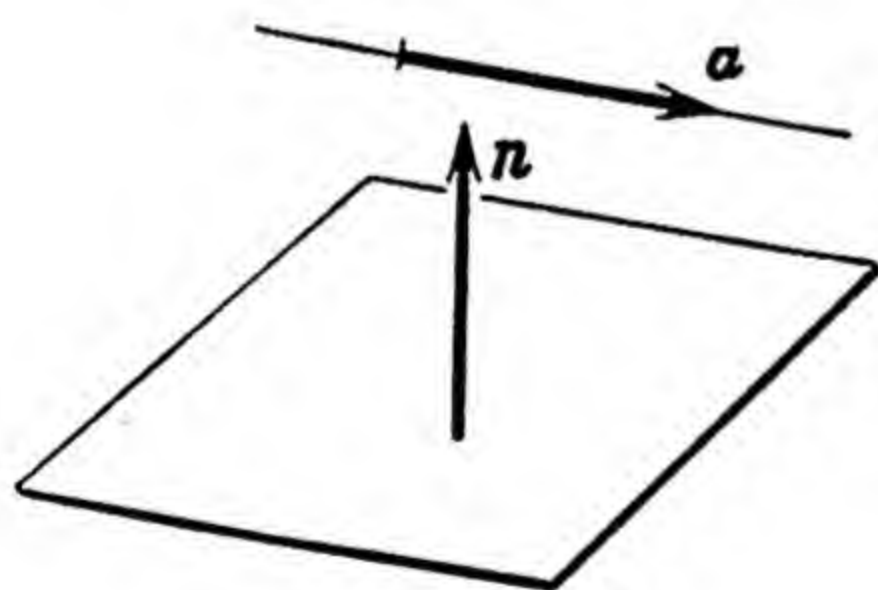


Fig. 110.

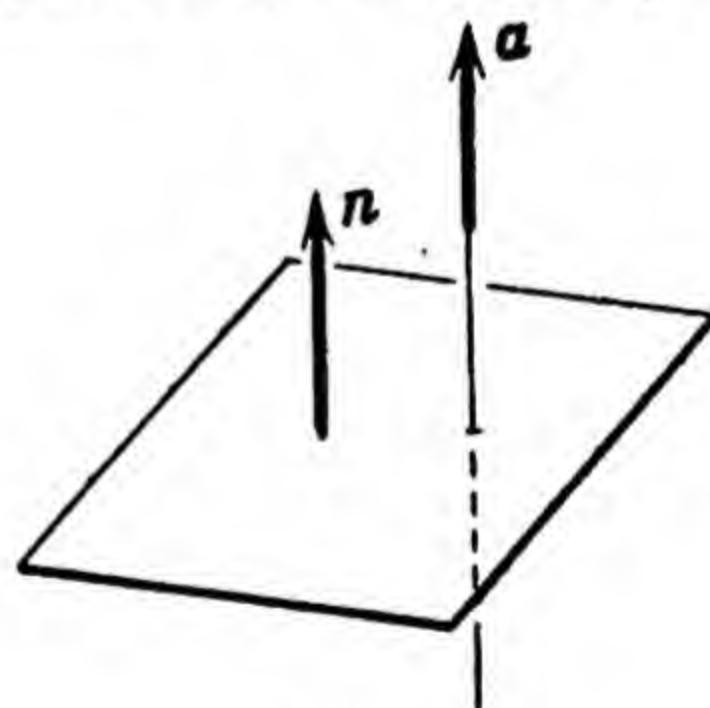


Fig. 111.

case of parallelism). Hence we have the following condition for the parallelism of a straight line and a plane:

$$Al + Bm + Cn = 0.$$

The straight line is perpendicular to the plane if, and only if, the direction vector of the line is collinear with the normal vector of the plane (Fig. 111). Hence we obtain the following condition

for the perpendicularity of a straight line and a plane:

$$\frac{A}{l} = \frac{B}{m} = \frac{C}{n}.$$

Below we give a few numerical examples.

221. Example 1. Find the equation of the plane passing through the line

$$\left. \begin{aligned} 3x + 2y + 5z + 6 &= 0, \\ x + 4y + 3z + 4 &= 0 \end{aligned} \right\}$$

and parallel to the line

$$\frac{x-1}{3} = \frac{y-5}{2} = \frac{z+1}{-3}.$$

Solution. Let us form the equation of the pencil of planes (see Arts 212, 213) passing through the first of the given lines:

$$3x + 2y + 5z + 6 + \lambda(x + 4y + 3z + 4) = 0. \quad (1)$$

From this pencil, we are to select the plane parallel to the second line; this amounts to finding the appropriate value of λ . Write equation (1) as

$$(3 + \lambda)x + (2 + 4\lambda)y + (5 + 3\lambda)z + (6 + 4\lambda) = 0. \quad (2)$$

The required plane must be parallel to the line

$$\frac{x-1}{3} = \frac{y-5}{2} = \frac{z+1}{-3}.$$

Using the parallelism condition for a line and a plane, we obtain the following equation in the unknown λ :

$$3(3 + \lambda) + 2(2 + 4\lambda) - 3(5 + 3\lambda) = 0.$$

Hence $\lambda = 1$. Substituting this value for λ in equation (2), we find: $4x + 6y + 8z + 10 = 0$, or $2x + 3y + 4z + 5 = 0$.

Example 2. Given the line

$$\left. \begin{aligned} 3x - 2y - z + 4 &= 0, \\ x - 4y - 3z - 2 &= 0; \end{aligned} \right\}$$

find its projection on the plane $5x + 2y + 2z - 7 = 0$.

Solution. We must find the plane passing through the given line and perpendicular to the given plane; the desired projection will then be determined as the intersection of this plane and the given plane. Let us form the equation of the pencil of planes passing through the given line:

$$3x - 2y - z + 4 + \lambda(x - 4y - 3z - 2) = 0. \quad (3)$$

This equation represents the required plane for a certain value of λ , which we shall now determine. Write equation (3) as

$$(3 + \lambda)x + (-2 - 4\lambda)y + (-1 - 3\lambda)z + (4 - 2\lambda) = 0. \quad (4)$$

The required plane must be perpendicular to the given plane. Using the perpendicularity condition for two planes, we get the following equation in the unknown λ :

$$5(3 + \lambda) + 2(-2 - 4\lambda) + 2(-1 - 3\lambda) = 0.$$

Hence $\lambda = 1$. Inserting the value of λ in (4), we find the equation of the plane passing through the given line and perpendicular to the given plane: $4x - 6y - 4z + 2 = 0$, or $2x - 3y - 2z + 1 = 0$. The projection of the given line on the given plane is thus represented by the equations

$$\begin{cases} 2x - 3y - 2z + 1 = 0, \\ 5x + 2y + 2z - 7 = 0. \end{cases}$$

Example 3. Find the distance from the point $P(1, 1, 1)$ to the line

$$\frac{x-11}{2} = \frac{y-18}{5} = \frac{z-4}{-2}.$$

Solution. Through P , pass a plane α perpendicular to the given line, and find the point Q in which this plane intersects the given line. The desired distance from the point P to the given line will be equal to the distance from the point P to the point Q .

By Art. 199, the equation of the plane α may be written in the form

$$A(x-1) + B(y-1) + C(z-1) = 0;$$

this plane must be perpendicular to the given line. By the perpendicularity condition for a line and a plane, we have

$$\frac{A}{2} = \frac{B}{5} = \frac{C}{-2};$$

letting, for simplicity, the factor of proportionality to be equal to unity, we find $A = 2$, $B = 5$, $C = -2$. Thus, the plane has as its equation $2(x-1) + 5(y-1) - 2(z-1) = 0$, or $2x + 5y - 2z - 5 = 0$. Next, we must find the point Q where this plane intersects the given line. This is achieved by solving the equations of the given line simultaneously with the equation of the plane α just found. Proceeding as shown in Art. 216 (see the example at the end of that article), we find the coordinates of the point Q to be $x = 5$, $y = 3$, $z = 10$. The desired distance d from the point P to the given line, which is equal to the distance between the points P and Q , is then found from the well-known distance formula:

$$d = \sqrt{(5-1)^2 + (3-1)^2 + (10-1)^2} = \sqrt{101} \approx 10.$$

QUADRIC SURFACES

§ 69. The Ellipsoid and the Hyperboloids

222. According to Art. 196, quadric surfaces (i. e., surfaces of the second order) are those represented by an equation of the second degree in cartesian coordinates. In this chapter we shall discuss various quadric surfaces. To begin with, we shall consider the ellipsoid and the two hyperboloids; these surfaces are the space analogues of plane ellipses and hyperbolas.

223. An *ellipsoid* is defined as the surface represented, in a rectangular cartesian system of coordinates, by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \quad (1)$$

Equation (1) is called *the canonical equation of an ellipsoid*.

In order to form a clear idea of the shape of an ellipsoid and to sketch the surface, we shall use the so-called "method of parallel sections".

Consider the sections of the given ellipsoid by planes parallel to the coordinate plane Oxy . Every such plane is represented by an equation of the form $z = h$, and the corresponding curve of intersection is represented by the two equations

$$\left. \begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} &= 1 - \frac{h^2}{c^2}, \\ z &= h. \end{aligned} \right\} \quad (2)$$

Hence it can be seen that: (1) when $|h| < c$, the plane $z = h$ intersects the ellipsoid in an ellipse having semi-axes

$$a^* = a \sqrt{1 - \frac{h^2}{c^2}},$$

$$b^* = b \sqrt{1 - \frac{h^2}{c^2}}$$

and symmetric with respect to the planes Oxz and Oyz ; (2) the quantities a^* and b^* have their largest values when $h = 0$ (then $a^* = a$, $b^* = b$); in other words, the largest of these ellipses is the section by the coordinate plane $z = 0$; (3) as $|h|$ increases,

the quantities a^* and b^* decrease; (4) when $h = \pm c$, the quantities a^* and b^* become zero, that is, the ellipse which is the section of the ellipsoid (1) by the plane $z = c$ or by the plane $z = -c$ degenerates into a point; in other words, the planes $z = \pm c$ are tangent planes to the ellipsoid; (5) when $|h| > c$, equations (2) represent an imaginary ellipse; this means that, when $|h| > c$, the plane $z = h$ does not meet the given ellipsoid at all.

We have an entirely analogous situation when considering the sections of the ellipsoid by planes parallel to the coordinate planes Oxz and Oyz . It will therefore be sufficient to note that the plane Oxz intersects the ellipsoid in the ellipse represented by the equations $\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1$ and $y = 0$, and that the plane Oyz intersects the ellipsoid in the ellipse represented by the equations $\frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ and $x = 0$ (see Fig. 112, showing the sections of the ellipsoid (1) by the planes Oxy , Oxz and $z = h$).

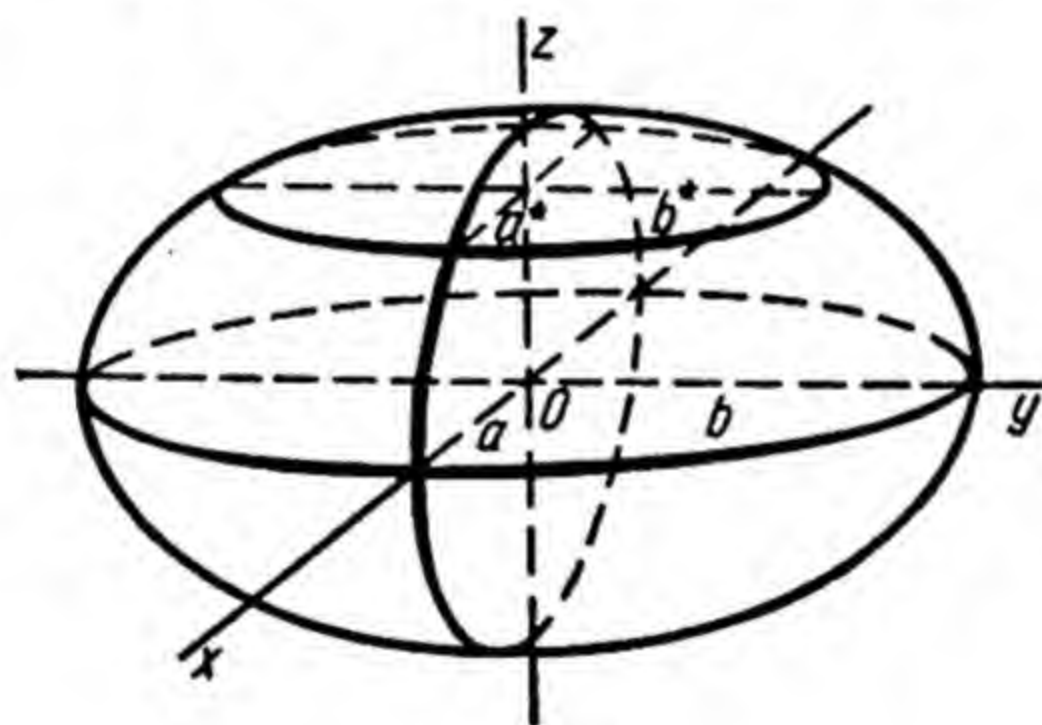


Fig. 112.

Bringing the above results together, we may conclude that an ellipsoid is a closed oval surface having three mutually perpendicular planes of symmetry. In the chosen coordinate system, these planes coincide with the coordinate planes.

224. The quantities a , b , c are called the semi-axes of an ellipsoid. If they are all of a different length, the ellipsoid is referred to as triaxial. Let us consider the case where two of the quantities a , b , c are equal. Let, for instance, $a = b$. Then equations (2) represent a circle with centre on the axis Oz . From this it follows that, when $a = b$, the ellipsoid may be thought of as the surface generated by revolving an ellipse about one of its axes. An ellipsoid generated by revolving an ellipse about its major axis, is called a *prolate ellipsoid of revolution*; an ellipsoid

obtained by revolving an ellipse about its minor axis is called an *oblate ellipsoid of revolution*. In the case $a = b = c$, the ellipsoid is a *sphere*.

225. Consider the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1. \quad (3)$$

Its left member is an expression identical with the left member of the canonical equation of an ellipsoid. Since this expression ≥ 0 , whereas the right member of (3) is -1 , it follows that equation (3) represents no real geometric object. In view of its similarity to (1), equation (3) is referred to as the equation of an *imaginary ellipsoid*.

226. We turn now to a consideration of the hyperboloids. There exist two of them: the hyperboloid of one sheet, and the hyperboloid of two sheets.

A *hyperboloid of one sheet* is the surface represented, in a rectangular cartesian system of coordinates, by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1. \quad (4)$$

A *hyperboloid of two sheets* is the surface represented by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1. \quad (5)$$

Equations (4) and (5) are called the canonical equations of the hyperboloids.

227. In this article we shall investigate the *hyperboloid of one sheet*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

Consider its sections by the coordinate planes Oxz and Oyz . The section by the plane Oxz is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{a^2} - \frac{z^2}{c^2} &= 1, \\ y &= 0. \end{aligned} \right\}$$

We see that this section is a hyperbola symmetric with respect to the coordinate axes Ox , Oz and cutting the axis Ox in the

points $(a, 0, 0)$ and $(-a, 0, 0)$. The section by the plane Oyz is represented by the equations

$$\left. \begin{aligned} \frac{y^2}{b^2} - \frac{z^2}{c^2} &= 1, \\ x &= 0; \end{aligned} \right\}$$

it is a hyperbola symmetric with respect to the axes Oy , Oz and cutting the axis Oy in the points $(0, b, 0)$ and $(0, -b, 0)$.

Consider now the sections of the given hyperboloid by planes parallel to the coordinate plane Oxy . Every such plane is represented by an equation of the form $z = h$, and the section of the hyperboloid by such a plane is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} &= 1 + \frac{h^2}{c^2}, \\ z &= h. \end{aligned} \right\} \quad (6)$$

Hence it is clear that: (1) every plane $z = h$ intersects the hyperboloid (4) in an ellipse having semi-axes

$$a^* = a \sqrt{1 + \frac{h^2}{c^2}}, \quad b^* = b \sqrt{1 + \frac{h^2}{c^2}}$$

and symmetric with respect to the planes Oxz and Oyz ; (2) the quantities a^* and b^* have their smallest values when $h = 0$ (then $a^* = a$, $b^* = b$); in other words, the smallest of these ellipses is the section by the coordinate plane $z = 0$ (this ellipse is called the *gorge* ellipse of a hyperboloid of one sheet); (3) as $|h|$ increases indefinitely, the quantities a^* and b^* also increase indefinitely (Fig. 113).

Summarising the above results, we may conclude that the hyperboloid of one sheet has the shape of an endless tube flaring out indefinitely on both sides of the gorge ellipse. The hyperboloid of one sheet possesses three mutually perpendicular planes of symmetry; in the coordinate system chosen, these planes coincide with the coordinate planes.

228. The quantities a , b , c are called *the semi-axes* of the hyperboloid of one sheet. The first two of them (a and b) are shown in Fig. 113. To draw the semi-axis c , it would be necessary to construct the fundamental rectangle of either of the two hyperbolas which are the sections of the hyperboloid of one sheet by the planes Oxz and Oyz .

Note that, in the case $a = b$, equations (6) represent a circle with centre on the axis Oz . It follows that, when $a = b$, the hyperboloid of one sheet may be thought of as the surface generated

by revolving a hyperbola about one of its axes, namely, about that axis which does not intersect the hyperbola.

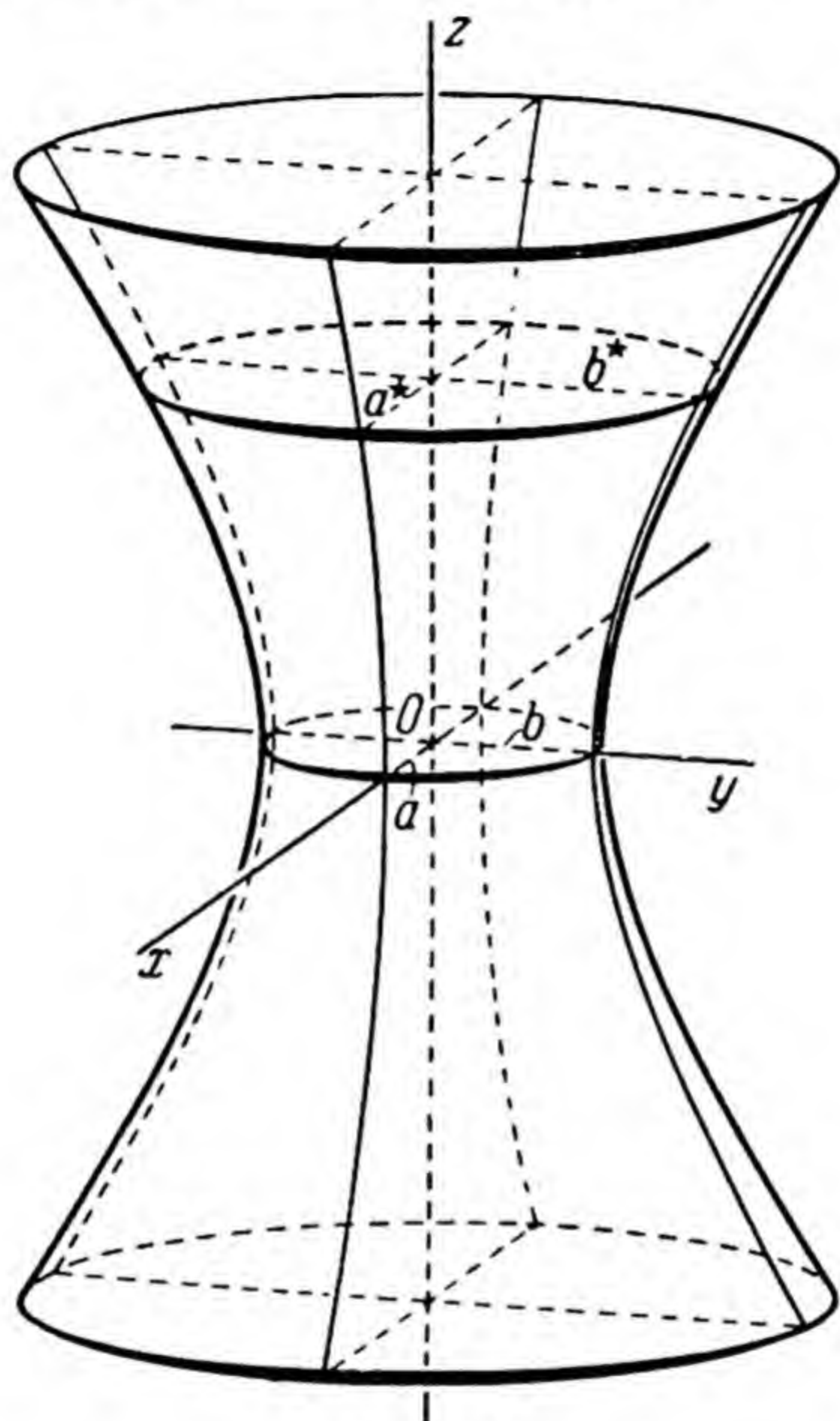


Fig. 113.

229. We now proceed to investigate *the hyperboloid of two sheets*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1.$$

Consider its sections by the coordinate planes Oxz and Oyz . The section by the plane Oxz is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{a^2} - \frac{z^2}{c^2} &= -1, \\ y &= 0. \end{aligned} \right\}$$

We see that this section is a hyperbola symmetric with respect to the coordinate axes Ox , Oz and cutting the axis Oz in the points $(0, 0, c)$ and $(0, 0, -c)$. The section by the plane Oyz is represented by the equations

$$\left. \begin{aligned} \frac{y^2}{b^2} - \frac{z^2}{c^2} &= -1, \\ x &= 0; \end{aligned} \right\}$$

it is a hyperbola symmetric with respect to the axes Oy , Oz and cutting the axis Oz (also in the points $(0, 0, c)$ and $(0, 0, -c)$).

Consider, finally, the sections of the given hyperboloid by planes parallel to the coordinate plane Oxy . Every such plane is represented by an equation of the form $z = h$, and the section of the hyperboloid by such a plane is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} &= \frac{h^2}{c^2} - 1, \\ z &= h. \end{aligned} \right\} \quad (7)$$

Hence we see that: (1) when $|h| > c$, the plane $z = h$ intersects the hyperboloid of two sheets in an ellipse having semi-axes

$$a^* = a \sqrt{\frac{h^2}{c^2} - 1},$$

$$b^* = b \sqrt{\frac{h^2}{c^2} - 1}$$

and symmetric with respect to the planes Oxz and Oyz ; (2) as $|h|$ increases, the quantities a^* and b^* increase; (3) as $|h|$ increases indefinitely, a^* and b^* do likewise; (4) as $|h|$ decreases and tends to c , a^* and b^* also decrease and tend to zero; for $h = \pm c$, we have $a^* = 0$, $b^* = 0$, which means that the ellipse constituting the

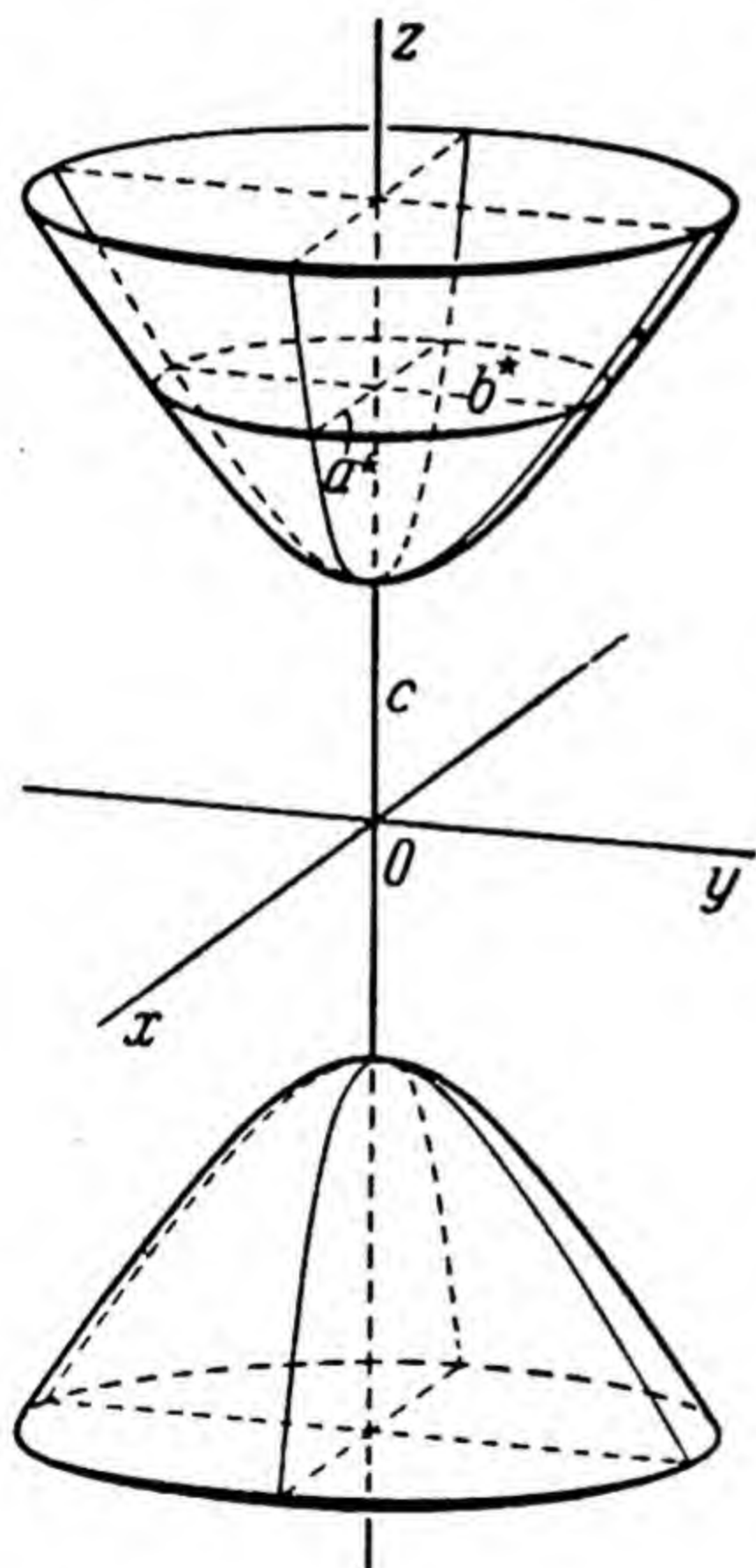


Fig. 114.

section by the plane $z = c$ or by the plane $z = -c$ degenerates into a point; in other terms, the planes $z = \pm c$ are tangent planes to the hyperboloid; (5) when $|h| < c$, equations (7) repre-

sent an imaginary ellipse; this means that, when $|h| < c$, the plane $z = h$ does not meet the given hyperboloid at all (Fig. 114).

Bringing together all these results, we conclude that a hyperboloid of two sheets is the surface composed of two distinct "sheets" (hence its name: "the hyperboloid of two sheets"), each of which has the shape of a convex bowl of endless extent. The hyperboloid of two sheets has three mutually perpendicular planes of symmetry; in the coordinate system chosen, these planes coincide with the coordinate planes.

230. The quantities a, b, c are called *the semi-axes* of the hyperboloid of two sheets. Shown in Fig. 114 is only the quantity c . To draw a and b , it would be necessary to construct the fundamental rectangles of the hyperbolas which are the sections of the hyperboloid of two sheets by the planes Oxz and Oyz .

Note that, in the case $a = b$, equations (7) represent a circle with centre on the axis Oz . It follows that, when $a = b$, the hyperboloid of two sheets may be regarded as the surface generated by revolving a hyperbola about one of its axes, namely, about the axis intersecting the hyperbola.

§ 70. The Quadric Cone

231. Consider the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0. \quad (1)$$

The distinctive feature of (1) is that it is a *homogeneous equation*, that is, all its terms are of the same degree ($=2$). Hence we have the following geometric property of the surface represented by this equation:

If a point M (other than the origin) lies on this surface, then all points of the straight line which passes through the origin and the point M also lie on this surface.

To prove our assertion, let M be a point with coordinates (l, m, n) , and let N be any point of the line OM . By Art. 216, the coordinates x, y, z of the point N are determined by the relations

$$x = lt, \quad y = mt, \quad z = nt,$$

where t is some number. Suppose that the point M lies on the surface under consideration; then

$$\frac{l^2}{a^2} + \frac{m^2}{b^2} - \frac{n^2}{c^2} = 0.$$

But in this case

$$\frac{(lt)^2}{a^2} + \frac{(mt)^2}{b^2} - \frac{(nt)^2}{c^2} = t^2 \left(\frac{l^2}{a^2} + \frac{m^2}{b^2} - \frac{n^2}{c^2} \right) = 0,$$

and hence the point N also lies on this surface. This completes the proof.

Note that this property is possessed by every surface which is represented in cartesian coordinates by a homogeneous equation (since the above argument was based solely on the homogeneity of a given equation). In other words, a surface represented

by a *homogeneous* equation is made up of straight lines all passing through the same point, namely, through the origin. Such a surface is called a *conic surface*, or simply a *cone*. The straight lines making up the cone are called *its elements*, and the point through which all the elements pass is referred to as *the vertex* of the cone.

In particular, the surface represented, in a cartesian system of coordinates, by an equation of the form (1) is called a *quadric cone*.

To form a clear idea of the shape of a quadric cone, we have merely to consider its section by a plane not passing through the origin (that is, by a plane not passing through the vertex of the cone). Take, for example, the plane $z = c$. The section of the cone by this plane is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} &= 1, \\ z &= c. \end{aligned} \right\} \quad (2)$$

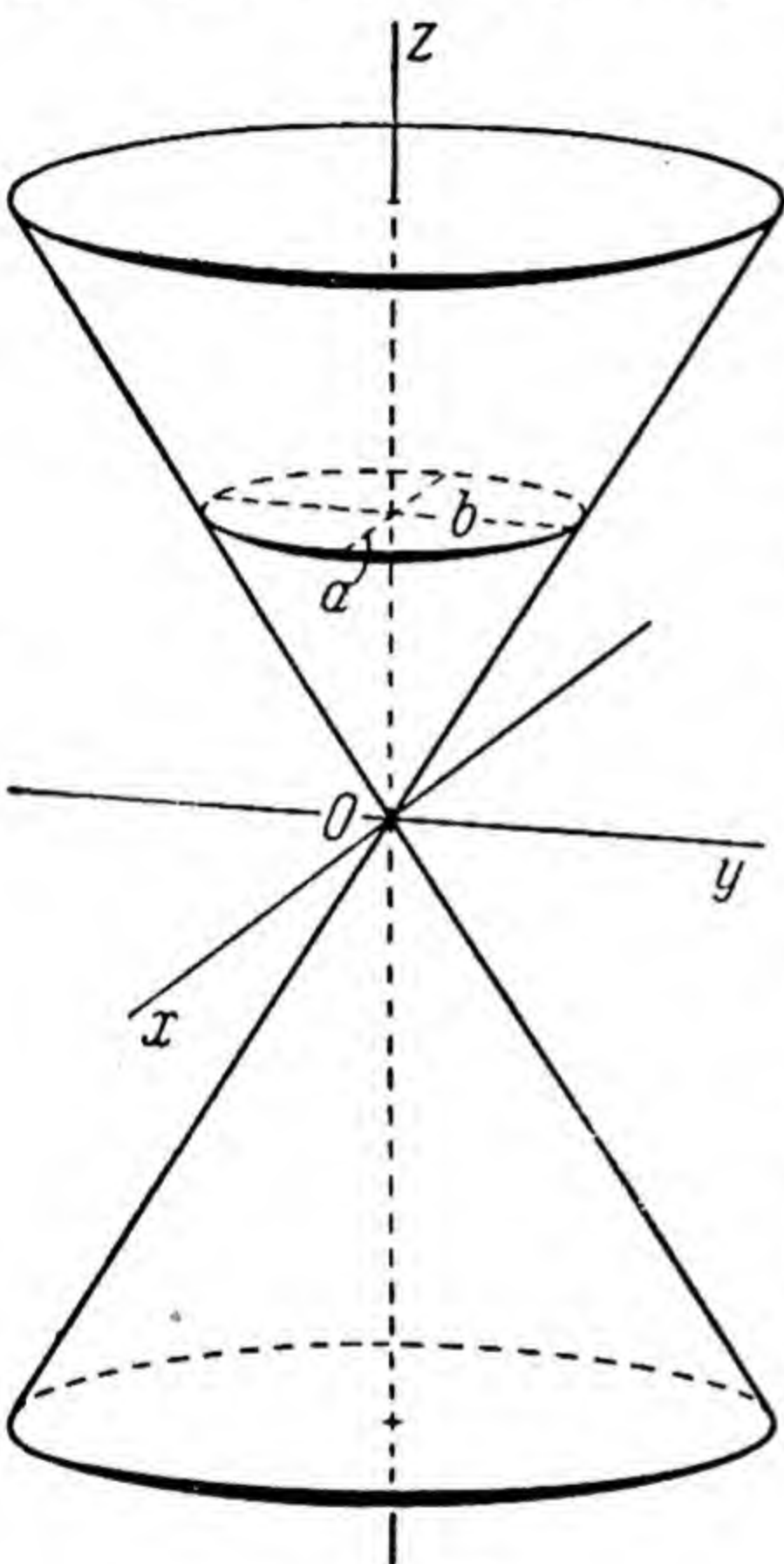


Fig. 115.

Obviously, this section is an ellipse having semi-axes a , b and symmetric with respect to the coordinate planes Oxz and Oyz .

The sketch of a quadric cone given in Fig. 115 has been drawn in accordance with this result.

Note that, if $a = b$, the ellipse represented by equations (2) becomes a circle with centre on the axis Oz and, consequently, the cone is then a *circular cone*.

232. Consider the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0; \quad (3)$$

it represents a single real point: $x = 0$, $y = 0$, $z = 0$. However, in view of its similarity to equation (1), equation (3) is often called the equation of an *imaginary cone*.

§ 71. The Paraboloids

233. There exist two surfaces which are the space analogues of plane parabolas. These surfaces are called the *paraboloids* (the elliptic paraboloid and the hyperbolic paraboloid).

234. An *elliptic paraboloid* is the surface which is represented, in a rectangular cartesian coordinate system, by the equation

$$2z = \frac{x^2}{p} + \frac{y^2}{q} \quad (1)$$

(where p and q are positive). Equation (1) is called the canonical equation of an elliptic paraboloid. We proceed to investigate this surface by the method of sections.

Let us consider, first of all, the sections by the coordinate planes Oxz and Oyz . Setting $y = 0$, we have, from (1), $x^2 = 2pz$; thus, the section by the plane Oxz is represented by the equations

$$\left. \begin{aligned} x^2 &= 2pz, \\ y &= 0. \end{aligned} \right\}$$

We see that the section is a parabola (with vertex at the origin) opening upwards and symmetric with respect to the axis Oz ; the parameter of this parabola is equal to p . The section by the plane Oyz is represented by the equations

$$\left. \begin{aligned} y^2 &= 2qz, \\ x &= 0 \end{aligned} \right\}$$

and is an analogously situated parabola with parameter q .

Let us now consider the sections of the given paraboloid by planes parallel to the coordinate plane Oxy . Every such plane is represented by an equation of the form $z = h$, and the section of the paraboloid by such a plane is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{p} + \frac{y^2}{q} &= 2h, \\ z &= h. \end{aligned} \right\} \quad (2)$$

Hence it can be seen that: (1) if $h > 0$, the plane $z = h$ intersects the elliptic paraboloid in an ellipse with semi-axes $a^* = \sqrt{2hp}$, $b^* = \sqrt{2hq}$ and symmetric with respect to the planes Oxz and Oyz ; (2) as h increases, the quantities a^* and b^* increase; (3) as h increases indefinitely, a^* and b^* do likewise; (4) as h decreases and tends to zero, a^* and b^* decrease and tend to zero, too; when $h = 0$, we have $a^* = 0$, $b^* = 0$; this means that the ellipse which is the section of the paraboloid (1) by the plane $z = 0$ degenerates into a point; in other terms, the plane $z = 0$ is tangent to the given elliptic paraboloid; (5) if $h < 0$, equations (2) represent an imaginary ellipse, which means that, if $h < 0$, the plane $z = h$ does not meet the given paraboloid at all (Fig. 116).

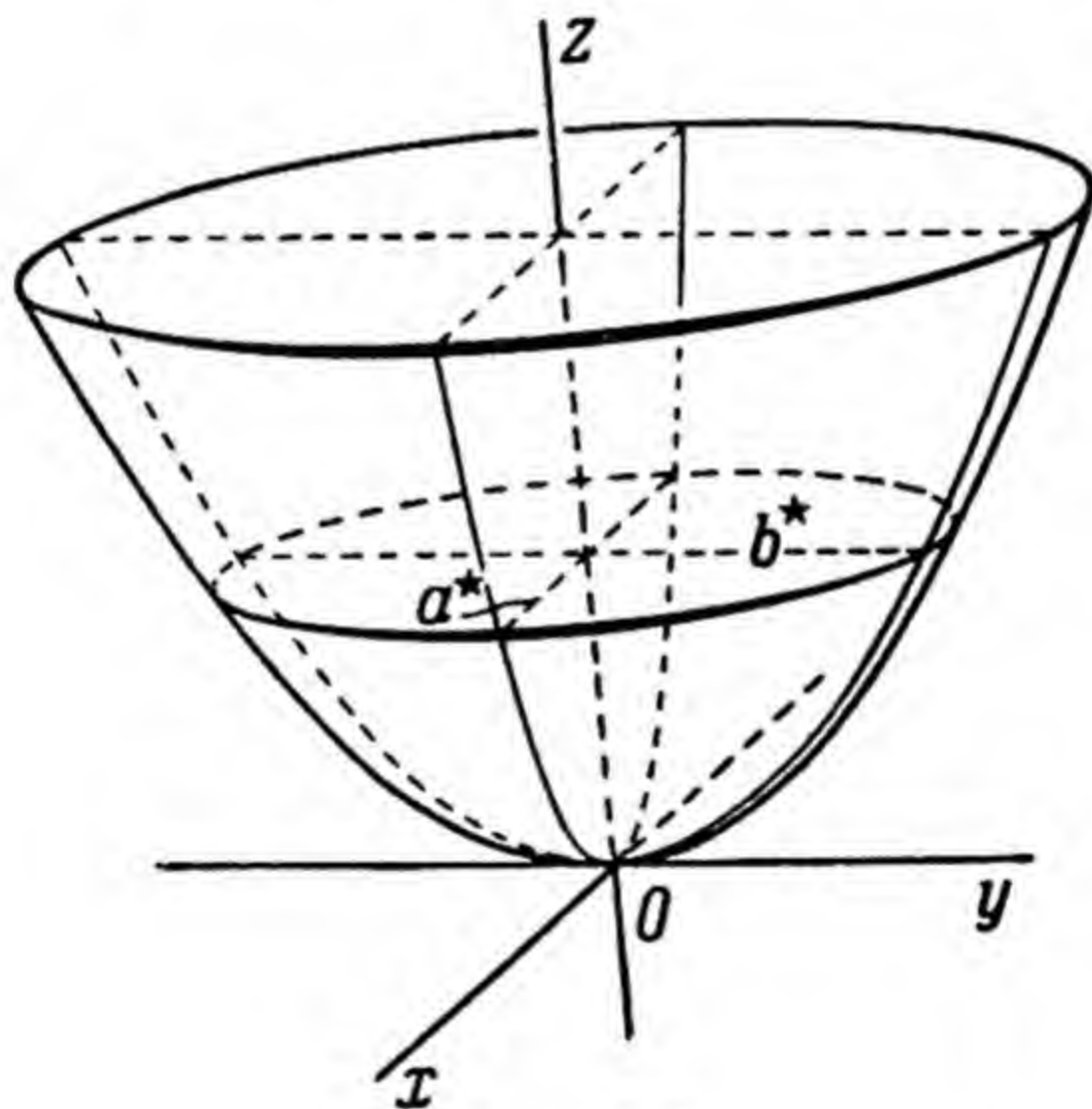


Fig. 116.

Summarising the above results, we conclude that an elliptic paraboloid has the shape of a convex bowl of endless extent. It possesses two mutually perpendicular planes of symmetry; in the chosen coordinate system, these planes coincide with the coordinate planes Oxz and Oyz . The point coincident with the origin of this system is called the *vertex* of the elliptic paraboloid; the numbers p and q are called the *parameters* of the elliptic paraboloid.

Note that, if $p = q$, equations (2) represent a circle with centre on the axis Oz . It follows that, if $p = q$, the elliptic paraboloid may be thought of as the surface generated by revolving a parabola about its axis.

235. The surface represented in a rectangular cartesian system of coordinates by an equation of the form

$$2z = \frac{x^2}{p} - \frac{y^2}{q} \quad (3)$$

(where p and q are positive) is called a hyperbolic paraboloid. Let us investigate this surface.

Consider the section of the hyperbolic paraboloid by the plane Oxz . Setting $y = 0$, we have, from (3), $x^2 = 2pz$; thus, the section by the plane Oxz is represented by the equations

$$\left. \begin{aligned} x^2 &= 2pz, \\ y &= 0. \end{aligned} \right\} \quad (4)$$

We see that this section is a parabola with vertex at the origin, opening upwards and symmetric with respect to the axis Oz ; the parameter of the parabola is equal to p .

Next consider the sections of the given paraboloid by planes parallel to the plane Oyz . Every such plane is represented by an equation of the form $x = h$, and the section of the paraboloid by such a plane is represented by the equations

$$\left. \begin{aligned} 2z &= -\frac{y^2}{q} + \frac{h^2}{p}, \\ x &= h. \end{aligned} \right\} \quad (5)$$

Hence we can see that, for any h , the plane $x = h$ intersects the hyperbolic paraboloid in a parabola open downwards and symmetric with respect to the plane Oxz (see Art. 120). It is apparent from the first of equations (5) that all these parabolas have the same parameter equal to q ; the vertex of each of these parabolas lies on the curve of intersection of the paraboloid and the plane Oxz (Fig. 117), that is, on the parabola open upwards and represented by equations (4).

Note that every plane $y = h$ intersects the hyperbolic paraboloid in a parabola open upwards, as is evident from the equations

$$\left. \begin{aligned} 2z &= \frac{x^2}{p} - \frac{h^2}{q}, \\ y &= h, \end{aligned} \right\}$$

which represent such sections; one of these sections, namely, that corresponding to $h = 0$, was considered at the beginning of the article.

A portion of a hyperbolic paraboloid is presented in Fig. 117; the edges of this portion are formed by two segments of parabolas opening upwards and lying in planes parallel to the plane Oxz ,

and by two segments of parabolas opening downwards and lying in planes parallel to the plane Oyz .

Finally, consider the sections of the hyperbolic paraboloid by planes parallel to the plane Oxy . Every such plane has $z = h$ as its equation, and the section of the paraboloid by such a plane is represented by the equations

$$\left. \begin{aligned} \frac{x^2}{p} - \frac{y^2}{q} &= 2h, \\ z &= h. \end{aligned} \right\}$$

Hence it is seen that the planes $z = h$ intersect the hyperbolic paraboloid in hyperbolas symmetric with respect to the planes

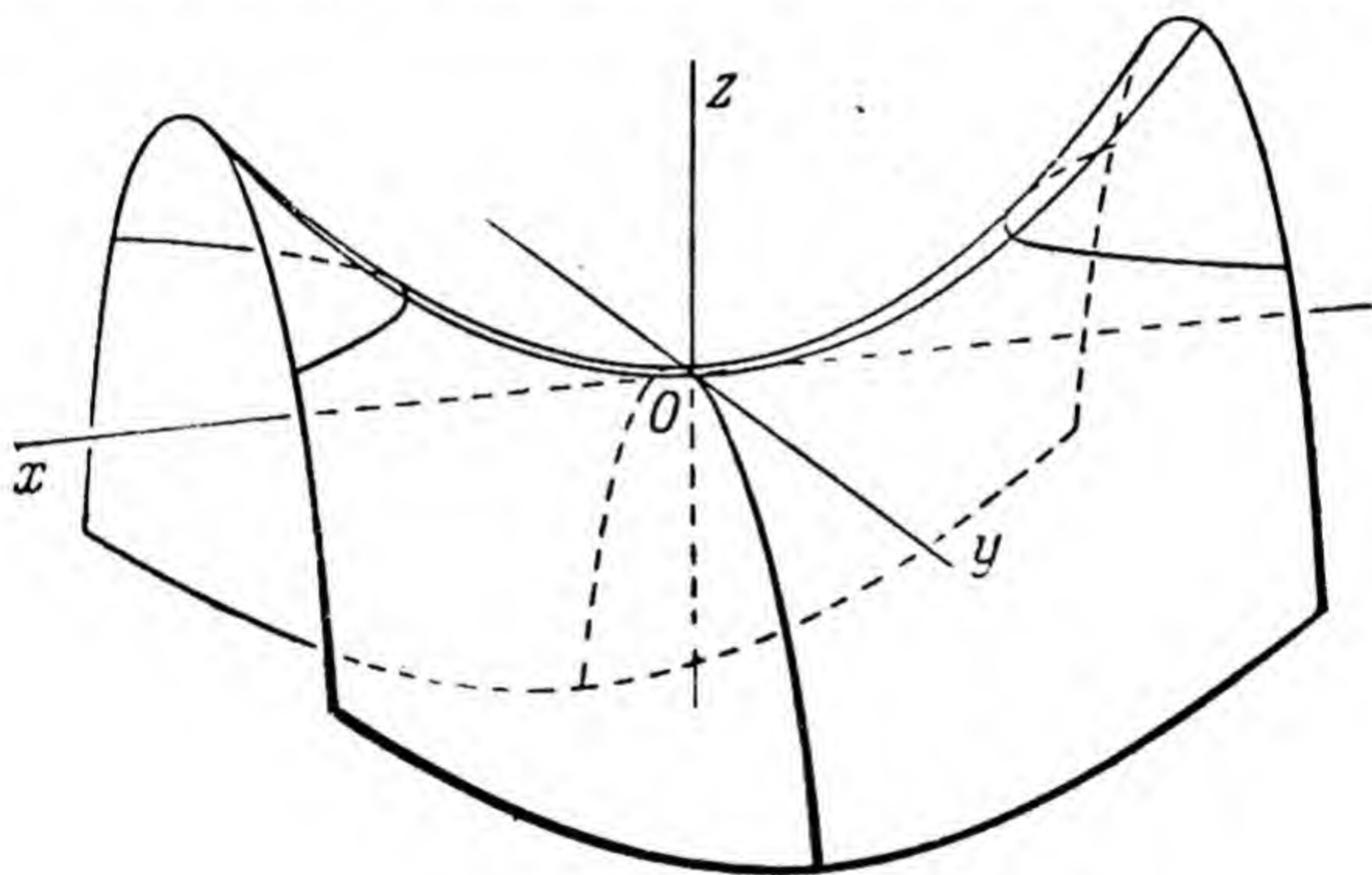


Fig. 117.

Oxz and Oyz . If $h > 0$, the corresponding hyperbolas intersect the plane Oxz ; if $h < 0$, the hyperbolas intersect the plane Oyz ; if $h = 0$, the corresponding hyperbola degenerates into a pair of lines. Fig. 117 shows the section of the paraboloid by one of the planes $z = h$ (for the case $h > 0$).

From all this, we may conclude that the hyperbolic paraboloid is saddle-shaped. It has two mutually perpendicular planes of symmetry; in the chosen coordinate system, these planes coincide with the coordinate planes Oxz and Oyz . The point coincident with the origin of this system is called *the vertex* of the hyperbolic paraboloid; the numbers p, q are called *the parameters* of the hyperbolic paraboloid.

§ 72. The Quadric Cylinders

236. To complete our study of quadric surfaces, let us consider a second-degree equation lacking the current coordinate z . We may write this equation in the form

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0. \quad (1)$$

By Art. 193, equation (1) represents a *cylindrical* surface (or more briefly, a cylinder) with elements parallel to the axis Oz .

Since (1) is an equation of the second degree, the surface represented by it is called a *quadric cylinder*.

Note now that equation (1) is in fact identical with equation (1) of § 41, which, referred to cartesian *plane* coordinates, repre-

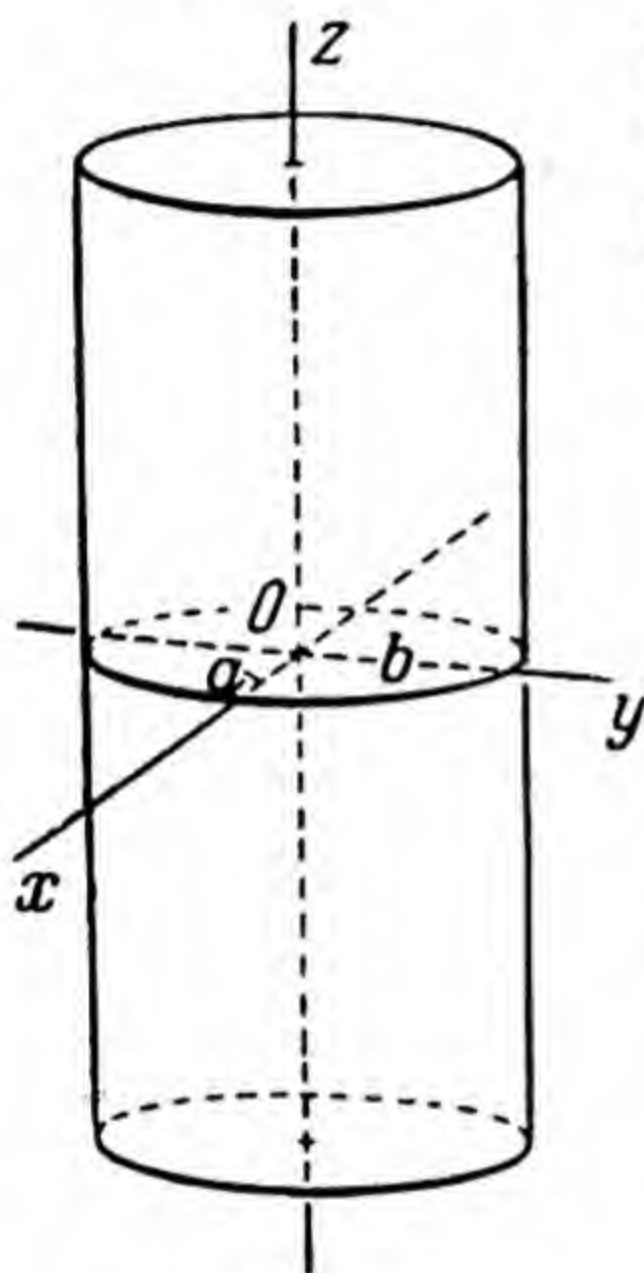


Fig. 118.

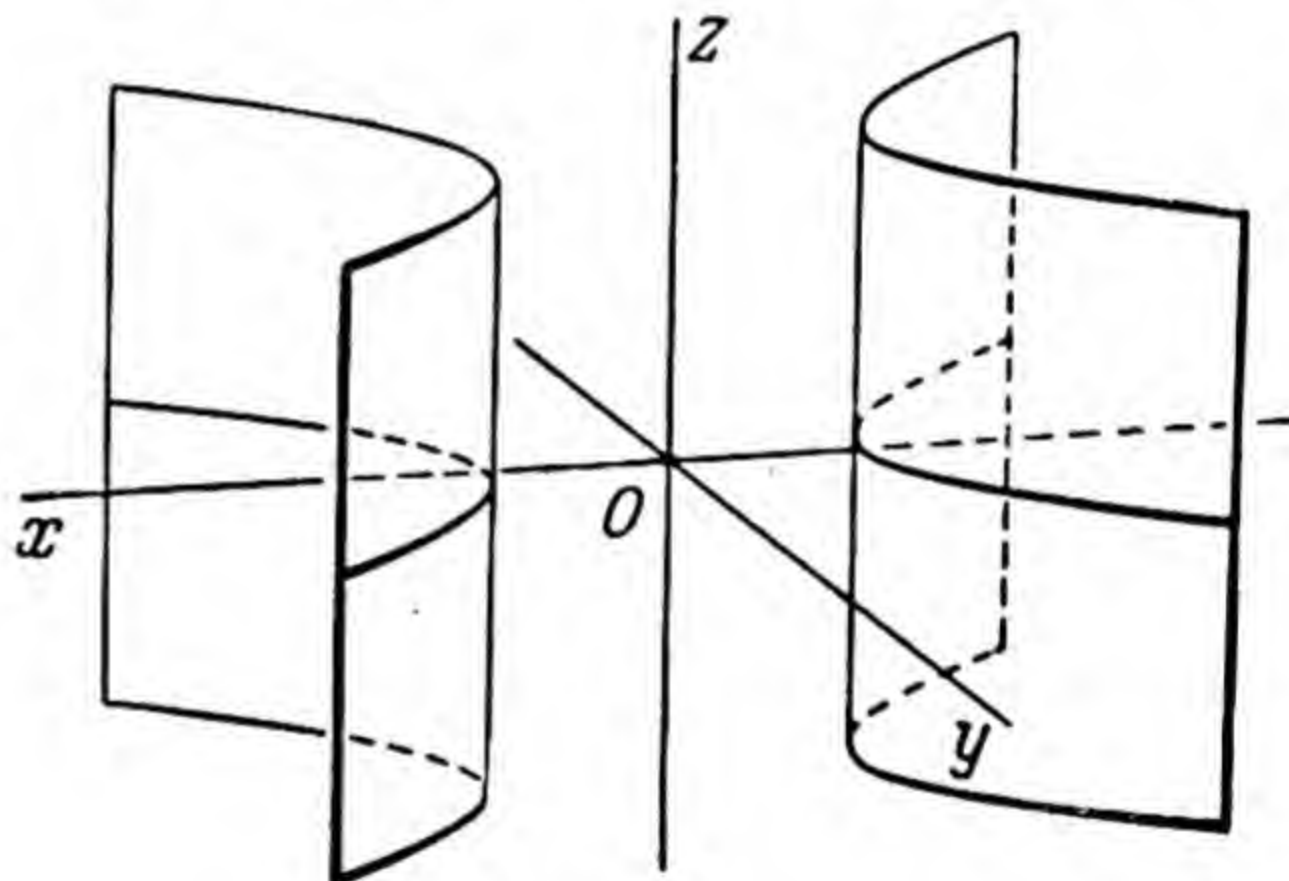


Fig. 119.

sents a curve of the second order. According to the nature of this curve, we have quadric cylinders of the following types:

(a) *The elliptic cylinder* (Fig. 118); by an appropriate choice of the coordinate system, its equation can be reduced to the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

If $a = b$, the cylinder is circular.

b) *The hyperbolic cylinder* (Fig. 119); its equation is reducible to the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1.$$

(c) *The parabolic cylinder* (Fig. 120); its equation is reducible to the form

$$y^2 = 2px.$$

Also, the left-hand member of (1) may happen to be the product of two first-degree factors. Then the cylinder “*degenerates*” into a pair of planes.

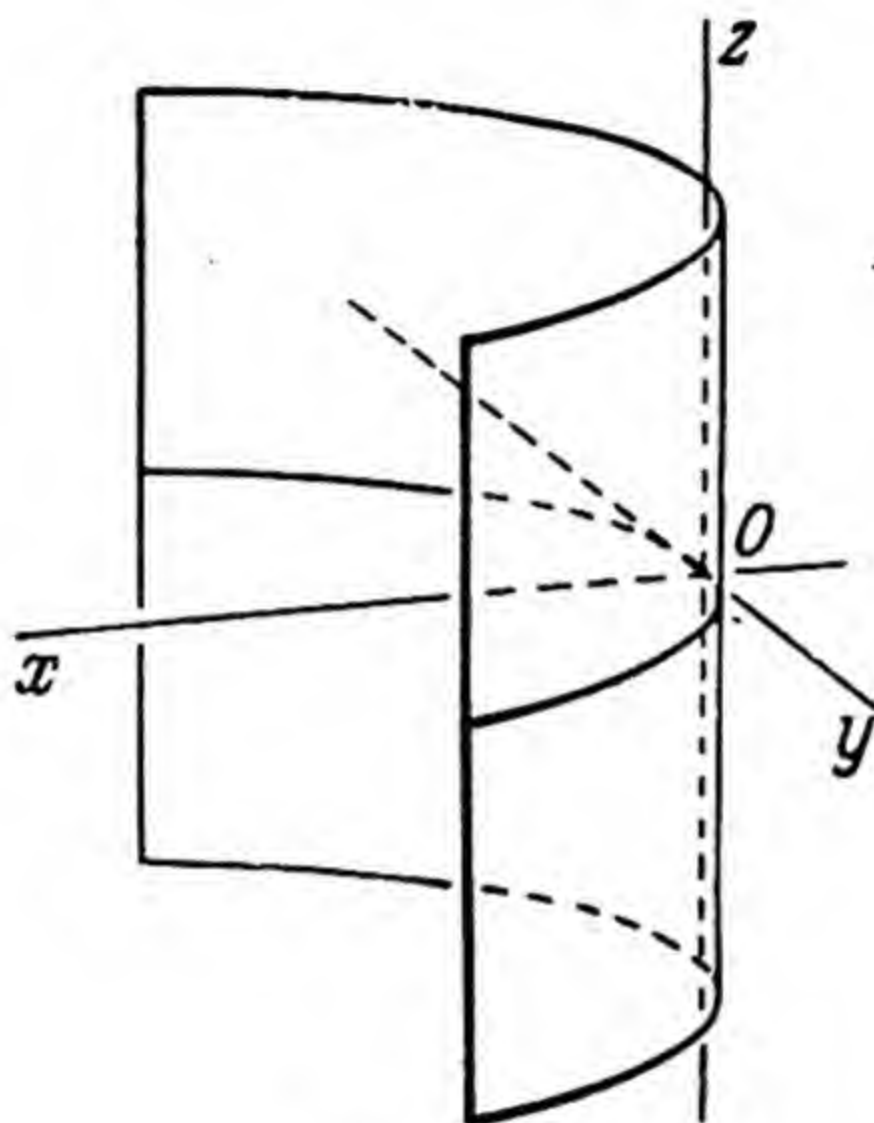


Fig. 120.

Finally, it may also occur that an equation of the form (1) has no real solutions at all (as for example, $x^2 + y^2 = -1$) and hence represents *no geometric object*. Such an equation is said to represent an imaginary cylinder.

§ 73. The Rectilinear Generators of the Hyperboloid of One Sheet. The Shukhov Towers

237. An inspection of the various types of quadric surfaces (see §§ 69-72) immediately reveals that some of them (namely, cones and cylinders) are *ruled* surfaces, that is, surfaces made up of straight lines. But, apart from cones and cylinders, the hyperboloid of one sheet and the hyperbolic paraboloid also turn out to be ruled surfaces. This fact is not revealed “by inspection”, but can readily be proved algebraically. We shall carry out the proof for the hyperboloid of one sheet.

Let the canonical equation of the hyperboloid of one sheet

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

be rewritten as

$$\frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2},$$

or

$$\left(\frac{x}{a} + \frac{z}{c}\right)\left(\frac{x}{a} - \frac{z}{c}\right) = \left(1 + \frac{y}{b}\right)\left(1 - \frac{y}{b}\right). \quad (1)$$

Consider next the two equations of the first degree

$$\left. \begin{aligned} \alpha\left(\frac{x}{a} + \frac{z}{c}\right) &= \beta\left(1 + \frac{y}{b}\right), \\ \beta\left(\frac{x}{a} - \frac{z}{c}\right) &= \alpha\left(1 - \frac{y}{b}\right), \end{aligned} \right\} \quad (2)$$

where α and β are some numbers, not both zero. When α and β have fixed values, equations (2) together represent a straight line; by varying the values of α and β , we obtain an infinite system of straight lines. Note now that equations (2), multiplied together, term by term, give equation (1). Hence *each of these straight lines lies entirely on the hyperboloid of one sheet*. For, if the coordinates x, y, z of a point satisfy both equations (2), they also satisfy equation (1); thus, every point of the line represented by equations (2), for any values of α, β (not both zero), lies on our hyperboloid of one sheet, which means that the entire line lies on the hyperboloid.

Finally, let us show that *through each point of the hyperboloid of one sheet there passes one, and only one, line of the system*. Let $M_0(x_0, y_0, z_0)$ be an arbitrary point of the hyperboloid of one sheet; since its coordinates satisfy the equation of the hyperboloid, it follows that

$$\left(\frac{x_0}{a} + \frac{z_0}{c}\right)\left(\frac{x_0}{a} - \frac{z_0}{c}\right) = \left(1 + \frac{y_0}{b}\right)\left(1 - \frac{y_0}{b}\right). \quad (3)$$

Let us find numbers α, β such that the corresponding line of the system (2) will pass through the point M_0 . Since the coordinates of M_0 must satisfy the equations of this line, we have the following two equations for the determination of the unknowns α, β :

$$\left. \begin{aligned} \alpha\left(\frac{x_0}{a} + \frac{z_0}{c}\right) &= \beta\left(1 + \frac{y_0}{b}\right), \\ \beta\left(\frac{x_0}{a} - \frac{z_0}{c}\right) &= \alpha\left(1 - \frac{y_0}{b}\right). \end{aligned} \right\} \quad (4)$$

If $1 + \frac{y_0}{b} \neq 0$, we find from the first equation of this system:

$$\beta = k\alpha,$$

where

$$k = \frac{\frac{x_0}{a} + \frac{z_0}{c}}{1 + \frac{y_0}{b}}. \quad (5)$$

When $\beta = k\alpha$, the second equation of the system (4) is also satisfied; this follows from relations (3) and (5). Substitute $\beta = k\alpha$ in equations (2), letting α have any value except zero. Since, after this substitution, both members of each of the equations contain the factor α , this factor can be divided out. We thus obtain a completely determined pair of equations,

$$\left. \begin{aligned} \frac{x}{a} + \frac{z}{c} &= k \left(1 + \frac{y}{b} \right), \\ k \left(\frac{x}{a} - \frac{z}{c} \right) &= 1 - \frac{y}{b}, \end{aligned} \right\}$$

to which there corresponds one completely determined straight line; this line passes through the point M_0 (since the numbers α and β have been chosen in conformity with relations (4)).

On the other hand, if $1 + \frac{y_0}{b} = 0$, then formula (5) has no meaning; but if $1 + \frac{y_0}{b} = 0$, it necessarily follows that $1 - \frac{y_0}{b} \neq 0$. Here the solution of the system (4) can be found from the second of its equations, after which it may be shown in a manner analogous to that used above that, also in this case, one and only one line of the system (2) passes through the point M_0 .

Thus, for various values of α and β , equations (2) represent an infinite system of straight lines (rulings), which lie on the hyperboloid of one sheet and entirely cover the surface. These rulings are called the *rectilinear generators* of the hyperboloid of one sheet.

We have shown that the hyperboloid of one sheet is made up of straight lines, i. e., that it is a *ruled* surface. Moreover, the hyperboloid of one sheet is a doubly ruled surface; this means that it has *two systems* of rectilinear generators.

For, analogous to equations (2), we can form the equations

$$\left. \begin{aligned} \alpha \left(\frac{x}{a} + \frac{z}{c} \right) &= \beta \left(1 - \frac{y}{b} \right), \\ \beta \left(\frac{x}{a} - \frac{z}{c} \right) &= \alpha \left(1 + \frac{y}{b} \right). \end{aligned} \right\} \quad (6)$$

Equations (6) also represent a system of rectilinear generators of the hyperboloid of one sheet, this system being different from that represented by equations (2).

A hyperboloid of one sheet with its two systems of rectilinear generators is shown in Fig. 121.

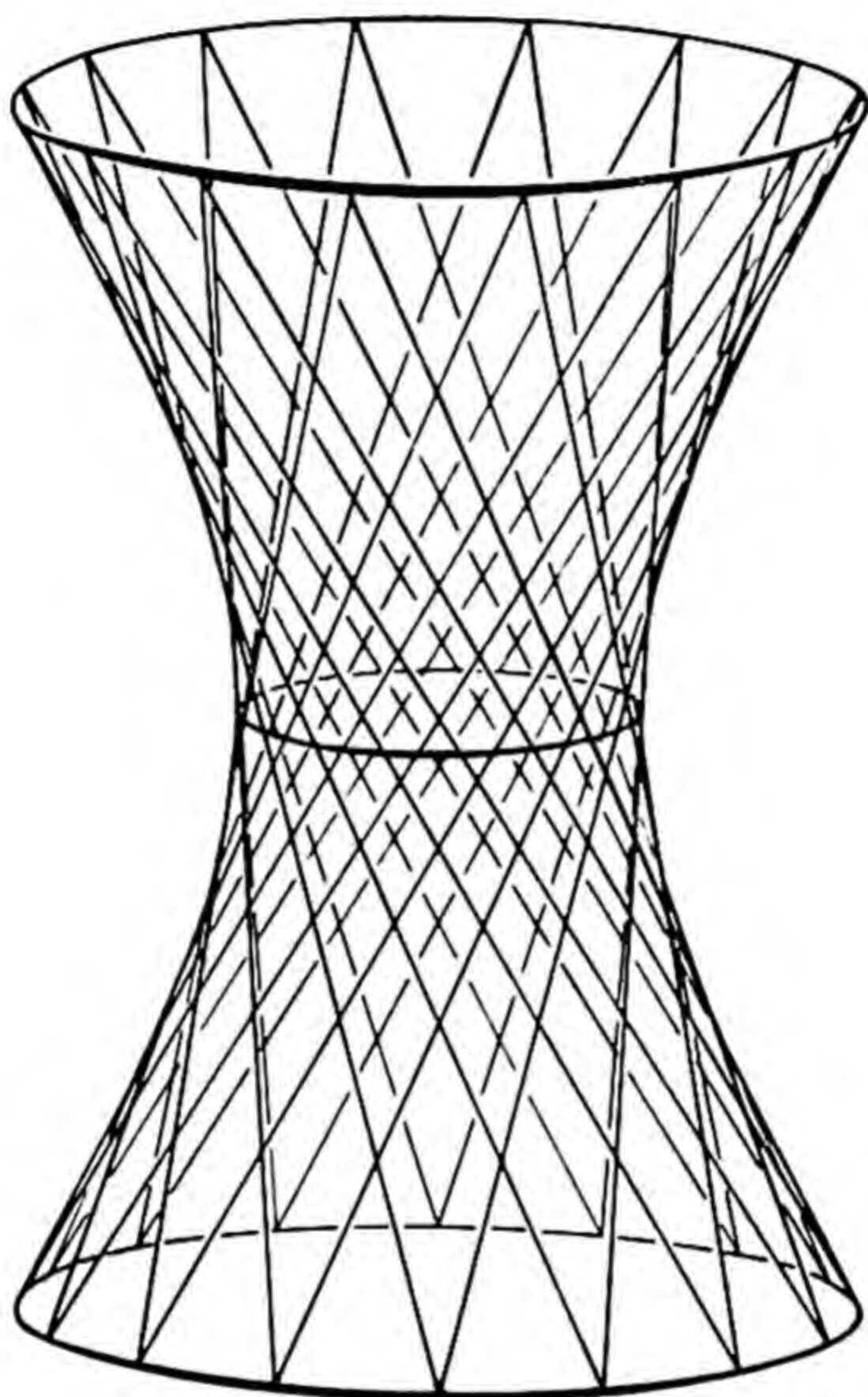


Fig. 121.

238. Without going into the details of the problem, let us point out that the *hyperbolic paraboloid*

$$2z = \frac{x^2}{p} - \frac{y^2}{q}$$

also has *two systems of rectilinear generators*; one of these systems is represented by the equations

$$\alpha \left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}} \right) = 2\beta z, \quad \beta \left(\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}} \right) = \alpha,$$

and the other system by the equations

$$\alpha\left(\frac{x}{\sqrt{p}} - \frac{y}{\sqrt{q}}\right) = 2\beta z, \quad \beta\left(\frac{x}{\sqrt{p}} + \frac{y}{\sqrt{q}}\right) = \alpha.$$

A hyperbolic paraboloid with its two systems of rectilinear generators is shown in Fig. 122.

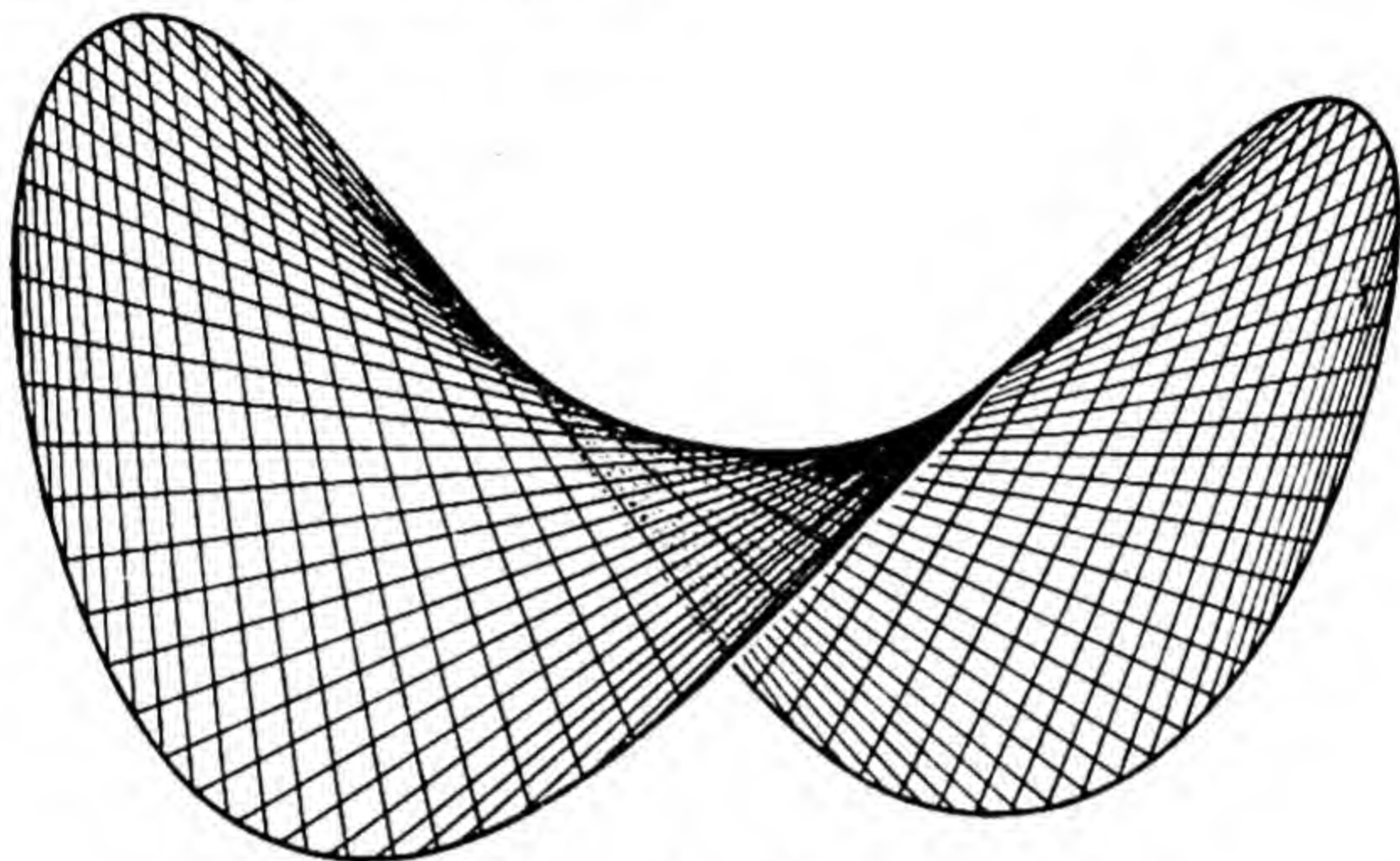


Fig. 122.

239. The idea of utilising the ruled nature of the hyperboloid of one sheet in constructional practice belongs to the famous Russian engineer Vladimir Shukhov. V. Shukhov invented structures from metal members arranged similar to the rectilinear generators of the hyperboloid of revolution of one sheet. These structures have proved to be both light and strong. They have often been used in the construction of water towers and high radio masts.

THE ELEMENTS OF THE THEORY OF DETERMINANTS

§ 1. Determinants of the Second Order and Systems of Two Equations of the First Degree in Two Unknowns

1. Consider a square array of four numbers, a_1, a_2, b_1, b_2 :

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}. \quad (1)$$

The number $a_1b_2 - a_2b_1$ is referred to as the determinant of the second order associated with the array (1). This determinant is denoted by the symbol $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$; accordingly,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1. \quad (2)$$

The numbers a_1, a_2, b_1, b_2 are called the elements of the determinant. The elements a_1, b_2 are said to lie on the principal diagonal of the determinant, and the elements a_2, b_1 , on the secondary diagonal. Thus, a determinant of the second order is equal to the product of the elements on the principal diagonal minus the product of the elements on the secondary diagonal.

For example,

$$\begin{vmatrix} -2 & 5 \\ -4 & 3 \end{vmatrix} = -2 \cdot 3 - (-4) \cdot 5 = 14.$$

2. We shall now show how determinants of the second order are used in analysing and solving a system of two equations of the first degree in two unknowns.

Consider the system of two equations

$$\left. \begin{aligned} a_1x + b_1y &= h_1, \\ a_2x + b_2y &= h_2 \end{aligned} \right\} \quad (3)$$

in the unknowns x and y (the coefficients a_1, b_1, a_2, b_2 and the constant terms h_1, h_2 are assumed to be known). The pair of numbers x_0, y_0 is said to be a solution of the system (3) if these numbers satisfy the system, that is, if each of equations (3) becomes an arithmetical identity after substituting the numbers x_0 and y_0 for x and y , respectively.

Let us find all solutions of the system (3); at the same time, let us analyse the system, namely, determine the cases where the system (3) has only one solution, more than one solution, and no solution at all. Using a well-known method of elimination (multiplying the first equation throughout by b_2 , the second by $-b_1$, and adding the results term by term), we eliminate the unknown y and obtain

$$(a_1b_2 - a_2b_1)x = b_2h_1 - b_1h_2. \quad (4)$$

Similarly, by eliminating the unknown x from the system (3), we find

$$(a_1b_2 - a_2b_1)y = a_1h_2 - a_2h_1. \quad (5)$$

Introducing the notation

$$\left. \begin{aligned} \Delta &= \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \\ \Delta_x &= \begin{vmatrix} h_1 & b_1 \\ h_2 & b_2 \end{vmatrix}, \\ \Delta_y &= \begin{vmatrix} a_1 & h_1 \\ a_2 & h_2 \end{vmatrix} \end{aligned} \right\} \quad (6)$$

we can write equations (4) and (5) as

$$\Delta \cdot x = \Delta_x, \quad \Delta \cdot y = \Delta_y. \quad (7)$$

The determinant Δ formed from the coefficients of the unknowns of the system (3) is called *the determinant of the system*. The determinant Δ_x is obtained by replacing the elements of the first column of Δ by the constant terms of the system (3); the determinant Δ_y is obtained from Δ by replacing the elements of the second column by the constant terms of (3).

Suppose that $\Delta \neq 0$; from equations (7), we then find

$$x = \frac{\Delta_x}{\Delta}, \quad y = \frac{\Delta_y}{\Delta},$$

or, in a fuller form,

$$x = \frac{\begin{vmatrix} h_1 & b_1 \\ h_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a_1 & h_1 \\ a_2 & h_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}. \quad (8)$$

Obviously, these formulas give the solution of the derived system consisting of equations (7). They also give the solution of the

original system (3). To verify the latter, the unknowns x, y in the left members of equations (3) must be replaced by their values from formulas (8); after this substitution (followed by expansion of the determinants $\Delta, \Delta_x, \Delta_y$ and by simplification easily performed by the reader), the left member of the first of equations (3) will be seen to equal the number h_1 , and the left member of the second equation to equal the number h_2 , which means that formulas (8) give the solution of the system (3).

Accordingly, we can formulate the following proposition: *If the determinant Δ of the system (3) is different from zero, the system has a unique solution determined by formulas (8).*

3. Suppose now that $\Delta = 0$. Then, if at least one of the determinants Δ_x, Δ_y is different from zero, the system (3) has no solutions at all (the equations of the system are said to be *inconsistent*).

For, if $\Delta = 0$, but at least one of the determinants Δ_x, Δ_y is not equal to zero, then at least one of equations (7) is impossible, which means that the system (7) has no solutions. But then the system (3) does not possess any solutions either, since the system (7) has been derived from the system (3), and so each solution of (3), if any such solution should exist, would also constitute a solution of the system (7).

On the other hand, if $\Delta = 0$ and also $\Delta_x = \Delta_y = 0$, the system (3) has infinitely many solutions (in this case, one equation of the system is a consequence of the other).

For, if $\Delta = \Delta_x = \Delta_y = 0$, that is, if

$$a_1b_2 - a_2b_1 = 0, \quad a_1h_2 - a_2h_1 = 0, \quad b_1h_2 - b_2h_1 = 0,$$

then the coefficients of the unknowns and the constant terms of the given equations are all in proportion. This means that one of the equations of the system can be obtained by multiplying the other equation throughout by a certain common factor, so that the system consists, essentially, of one equation, say $a_1x + b_1y = h_1$, the other equation being its consequence. But an equation of the form $a_1x + b_1y = h_1$ always has infinitely many solutions, since we can assign arbitrary values to one of the two unknowns x, y and find the corresponding values of the other unknown from the equation (for example, if $b_1 \neq 0$, we can assign arbitrary values to x and determine y from the formula $y = \frac{-a_1x + h_1}{b_1}$).

Note. Our argument is based on the assumption that each separately taken equation of the system has a solution. If we include into consideration systems which contain contradictory equations, then the above proposition will no longer be true. For

instance, the system

$$\begin{aligned} 0 \cdot x + 0 \cdot y &= 1, \\ 0 \cdot x + 0 \cdot y &= 1 \end{aligned}$$

satisfies the conditions $\Delta = 0$, $\Delta_x = 0$, $\Delta_y = 0$; however, this system admits no solution.

4. To summarise, if the determinant of the system (3) is different from zero ($\Delta \neq 0$), the system has a unique solution given by formulas (8); if $\Delta = 0$, the system has either no solution at all, or infinitely many solutions.

Example 1. Find all solutions of the system

$$\begin{cases} 3x + 4y = 2, \\ 2x + 3y = 7. \end{cases}$$

Solution. We begin by evaluating the determinant of the system:

$$\Delta = \begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix} = 3 \cdot 3 - 2 \cdot 4 = 1.$$

Since $\Delta \neq 0$, the system has a unique solution determined by formulas (8). We now find Δ_x and Δ_y :

$$\Delta_x = \begin{vmatrix} 2 & 4 \\ 7 & 3 \end{vmatrix} = 2 \cdot 3 - 4 \cdot 7 = -22,$$

$$\Delta_y = \begin{vmatrix} 3 & 2 \\ 2 & 7 \end{vmatrix} = 3 \cdot 7 - 2 \cdot 2 = 17.$$

Hence

$$x = \frac{\Delta_x}{\Delta} = \frac{-22}{1} = -22, \quad y = \frac{\Delta_y}{\Delta} = \frac{17}{1} = 17.$$

Example 2. Find all solutions of the system

$$\begin{cases} 3x + 4y = 1, \\ 6x + 8y = 3. \end{cases}$$

Solution. The value of the determinant of the system is

$$\Delta = \begin{vmatrix} 3 & 4 \\ 6 & 8 \end{vmatrix} = 3 \cdot 8 - 4 \cdot 6 = 0.$$

Since $\Delta = 0$, the given system has either no solution, or an infinite number of solutions. To see which of the two possibilities materialises in the present case, we find Δ_x and Δ_y :

$$\Delta_x = \begin{vmatrix} 1 & 4 \\ 3 & 8 \end{vmatrix} = 8 - 12 = -4.$$

$$\Delta_y = \begin{vmatrix} 3 & 1 \\ 6 & 3 \end{vmatrix} = 9 - 6 = 3.$$

Since $\Delta = 0$, but $\Delta_x \neq 0$, $\Delta_y \neq 0$, our system has no solutions.

Note. The same conclusion can be reached at once by multiplying the first equation throughout by 2 and subtracting the result, term by term, from the second equation, which will give $0 = 1$, that is, a contradictory equation. Hence the given equations are inconsistent.

Example 3. Find all solutions of the system

$$\begin{cases} 3x + 4y = 1, \\ 6x + 8y = 2. \end{cases}$$

Solution. The coefficients of x and y are the same as in example 2; accordingly, $\Delta = 0$. Hence, our system has either no solution, or infinitely many solutions. But, as is easily seen, the second equation of the system is derivable from the first equation (by multiplying all terms of the latter by 2). Thus, the system reduces to a single equation and has, therefore, an infinite number of different solutions, which are found by assigning arbitrary values to x and determining the corresponding values of y from the formula

$$y = \frac{1 - 3x}{4}.$$

5. Consider, in particular, a system of two homogeneous equations in two unknowns,

$$\begin{cases} a_1x + b_1y = 0, \\ a_2x + b_2y = 0, \end{cases} \quad (9)$$

that is, a system of equations whose constant terms are all equal to zero.

Obviously, such a system always possesses *the zero solution* $x = 0, y = 0$. If $\Delta \neq 0$, this solution is unique; but if $\Delta = 0$, the homogeneous system has, in addition, an infinite number of solutions other than the zero solution (since, for a homogeneous system, the possibility of there being no solution at all is ruled out). This can also be formulated as follows: *The homogeneous system (9) has a non-zero solution if, and only if, $\Delta = 0$.*

§ 2. A Homogeneous System of Two Equations of the First Degree in Three Unknowns

6. Let us solve the system of two homogeneous equations

$$\begin{cases} a_1x + b_1y + c_1z = 0, \\ a_2x + b_2y + c_2z = 0 \end{cases} \quad (1)$$

in the three unknowns x, y, z . Suppose that

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0. \quad (2)$$

Rewrite the system (1) in the form

$$\begin{cases} a_1x + b_1y = -c_1z, \\ a_2x + b_2y = -c_2z \end{cases} \quad (3)$$

and assume that some arbitrary value has been assigned here to the unknown z . For a definite value of z , the system (3) has a unique solution, which is obtained by applying formulas (8) of § 1:

$$\left. \begin{aligned} x &= \frac{\begin{vmatrix} -c_1 z & b_1 \\ -c_2 z & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \\ y &= \frac{\begin{vmatrix} a_1 & -c_1 z \\ a_2 & -c_2 z \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}. \end{aligned} \right\} \quad (4)$$

The numbers x , y together with the number z constitute a solution of the given system (1); to different values of z there correspond different solutions of the system (1), which has an infinite number of solutions (since z may be chosen at will).

Let us give formulas (4) a more convenient form. First of all, note that

$$\begin{vmatrix} -c_1 z & b_1 \\ -c_2 z & b_2 \end{vmatrix} = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} z, \quad \begin{vmatrix} a_1 & -c_1 z \\ a_2 & -c_2 z \end{vmatrix} = - \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} z;$$

hence, formulas (4) may be written as

$$x = \frac{\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} z}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \quad y = \frac{- \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} z}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}. \quad (5)$$

Employing the notation

$$\Delta_1 = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad (6)$$

we rewrite formulas (5) in the form

$$x = \frac{\Delta_1 \cdot z}{\Delta_3}, \quad y = \frac{-\Delta_2 \cdot z}{\Delta_3}. \quad (7)$$

Denote $\frac{z}{\Delta_3}$ by the letter t ; then $z = \Delta_3 \cdot t$, and, according to formulas (7), x and y are expressed by the relations $x = \Delta_1 \cdot t$, $y = -\Delta_2 \cdot t$.

We thus get the formulas

$$x = \Delta_1 \cdot t, \quad y = -\Delta_2 \cdot t, \quad z = \Delta_3 \cdot t, \quad (8)$$

which determine all solutions of the system (1), each separate solution being obtained by assigning some definite value to t .

For practical calculations it will be helpful to observe that the determinants Δ_1 , Δ_2 , Δ_3 are obtained by deleting, in turn, each column of the array

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}.$$

7. The above derivation was based on the supposition that $\Delta_3 \neq 0$ (see relation (2)).

If $\Delta_3 = 0$, but at least one of the determinants Δ_1 , Δ_2 is not equal to zero, the derivation remains the same, except that the unknowns interchange roles (for example, if $\Delta_2 \neq 0$, then we assume that y is assigned arbitrary values, and determine the corresponding values of x and z from the equations of the system). The ultimate result is the same, that is, all solutions of the system are again determined by formulas (8).

If the three determinants Δ_1 , Δ_2 , Δ_3 are all equal to zero, that is, if

$$b_1c_2 - b_2c_1 = 0, \quad a_1c_2 - a_2c_1 = 0, \quad a_1b_2 - b_1a_2 = 0,$$

then the coefficients of the equations (1) are all in proportion, which means that one equation of the system is a consequence of the other: one equation can be obtained by multiplying all terms of the other by some numerical factor. Thus, if $\Delta_1 = 0$, $\Delta_2 = 0$, $\Delta_3 = 0$, the system reduces in fact to a single equation. Such a system naturally has an infinite number of solutions; to get one of these, it is necessary to assign arbitrary values to *two unknowns* and find the third unknown from the equation.

Example 1. Find all solutions of the system

$$\left. \begin{aligned} 3x + 5y + 8z &= 0, \\ 7x + 2y + 4z &= 0. \end{aligned} \right\}$$

Solution. By Art. 6 we have

$$\Delta_1 = 4, \quad \Delta_2 = -44, \quad \Delta_3 = -29.$$

All solutions of the given system are determined from the formulas

$$x = 4t, \quad y = 44t, \quad z = -29t,$$

where t may assume any values.

Example 2. Find all solutions of the system

$$\left. \begin{aligned} 3x + 2y - 3z &= 0, \\ 6x + 4y - 6z &= 0. \end{aligned} \right\}$$

Solution. We have $\Delta_1 = 0$, $\Delta_2 = 0$, $\Delta_3 = 0$; the system contains, essentially, a single equation (the second equation being obtained by multiplying the first by 2). Any solution of the system consists of three numbers x , y , z , where x , y may be chosen at will and $z = \frac{3x + 2y}{3}$.

§ 3. Determinants of the Third Order

8. Consider a square array of nine numbers, $a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3$:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}. \quad (1)$$

The determinant of the third order associated with the array (1) is the number denoted by the symbol

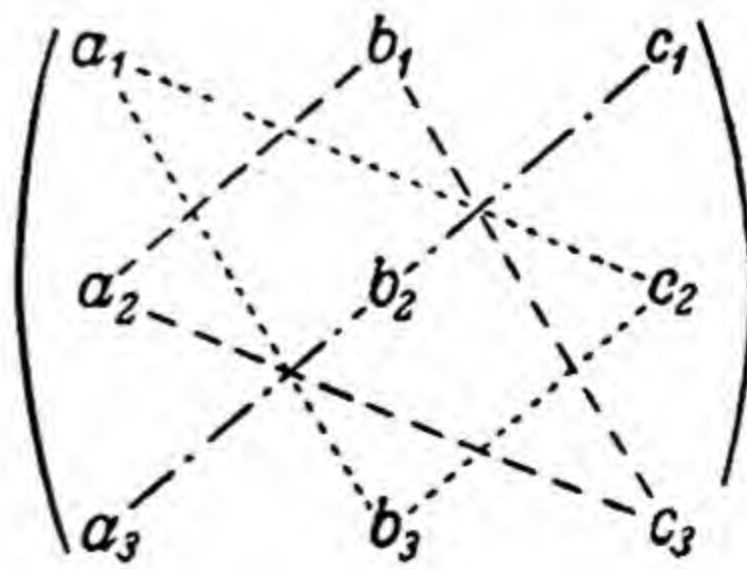
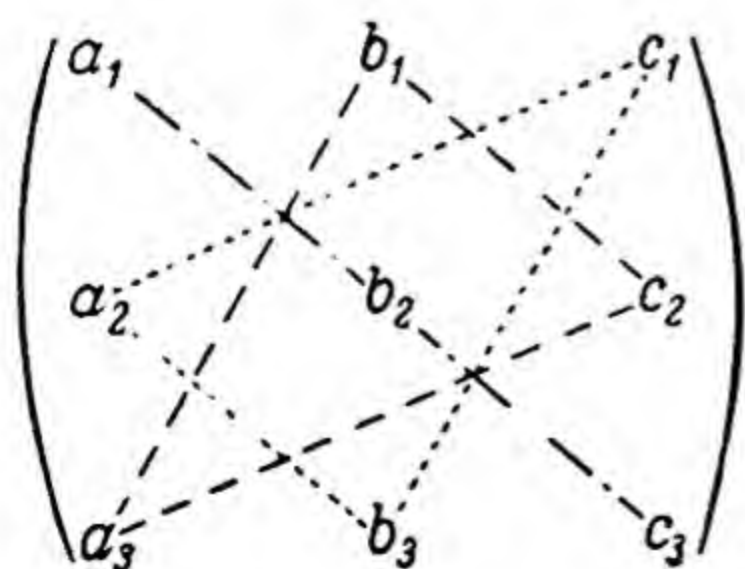
$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

and determined by the relation

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 + b_1 c_2 a_3 + c_1 a_2 b_3 - c_1 b_2 a_3 - b_1 a_2 c_3 - a_1 c_2 b_3. \quad (2)$$

The numbers $a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3$ are called *the elements of the determinant*. The diagonal containing the elements a_1, b_2, c_3 is called *the principal diagonal* of the determinant; the elements a_3, b_2, c_1 form the *secondary diagonal*.

The reader should observe that the first three terms in the right-hand member of (2) are the products of the elements taken three at a time as shown by the various dashed and dotted lines in the left-hand diagram below.



The remaining three terms are obtained by multiplying the elements three at a time as shown by the various lines in the right-hand diagram, and then changing the sign of each product.

This mnemonic rule, called the rule of triangles, not only facilitates writing out formula (2), but also permits us to evaluate a third-order determinant with numerical elements, without having first to write formula (2).

Thus, for example,

$$\begin{vmatrix} 3 & -2 & 1 \\ -2 & 1 & 3 \\ 2 & 0 & -2 \end{vmatrix} = 3 \cdot 1 \cdot (-2) + (-2) \cdot 3 \cdot 2 + (-2) \cdot 0 \cdot 1 - \\ - 2 \cdot 1 \cdot 1 - 3 \cdot 0 \cdot 3 - (-2) \cdot (-2) \cdot (-2) = -12.$$

9. Determinants are widely used both in pure and applied mathematics. We shall presently show how determinants of the third order are employed in analysing and solving a system of three first-degree equations in three unknowns. But we must first become familiar with certain properties of determinants. Some of their more important properties are considered in the next article, with determinants of the third order serving as illustrations throughout; the properties discussed are, however, equally valid for determinants of any order (the concept of a determinant of an order higher than the third is treated in the last section of this Appendix).

10. Property 1. *The value of a determinant is unchanged if all its columns are changed into rows so that each row is replaced by the like-numbered column, that is,*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \quad (3)$$

The same property may also be formulated as follows: *If the elements symmetrical with respect to the principal diagonal of a determinant are interchanged, the value of the determinant remains unchanged.*

To prove this, we have merely to expand both the left and the right member of (3) by using the rule of triangles, and to compare the results.

Note. The property 1 means the interchangeability of the rows and columns of a determinant; it will therefore be sufficient to prove the validity of the properties that follow (and that apply equally to rows and columns) for the rows or the columns only.

Property 2. *The interchange of two columns or two rows of a determinant is equivalent to multiplying the determinant by -1 .*

For example,

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = - \begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix}. \quad (4)$$

To prove relation (4), apply the rule of triangles to both the left and the right member of (4) and compare the results (the proof of analogous relations corresponding to the interchange of other columns is carried out in the same way).

Property 3. *If a determinant has two identical columns or two identical rows, the value of the determinant is zero.*

Let Δ be a determinant having two identical columns. If these columns are interchanged, the determinant changes its sign in virtue of the property 2. On the other hand, since the interchanged columns are identical, the interchange cannot alter the value of the determinant. Hence $\Delta = -\Delta$, that is, $2\Delta = 0$, or $\Delta = 0$.

For example,

$$\begin{vmatrix} 3 & 3 & 17 \\ 5 & 5 & 8 \\ 7 & 7 & 9 \end{vmatrix} = 0.$$

Property 4. *Multiplying all elements of a column or row by any one number k is equivalent to multiplying the determinant by this number k .*

This property may also be phrased as follows: A factor common to each element of a column or row can be taken out of that column or row and prefixed to the determinant.

For example,

$$\begin{vmatrix} ka_1 & b_1 & c_1 \\ ka_2 & b_2 & c_2 \\ ka_3 & b_3 & c_3 \end{vmatrix} = k \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

This is proved by simply observing that the expansion of a determinant is the sum of terms, each of which contains as a factor one element from each row and each column (see formula (2) of Art. 8).

Property 5. *If all elements of a column or row are zero, the determinant is zero.*

This property constitutes a special case (in which $k = 0$) of the preceding property.

For example,

$$\begin{vmatrix} 1 & 0 & 5 \\ 3 & 0 & 8 \\ 7 & 0 & 2 \end{vmatrix} = 0.$$

Property 6. *If the corresponding elements of two columns (or two rows) of a determinant are proportional, the determinant is zero.*

This follows from the properties 4 and 3. For, if the elements of two columns of a determinant are proportional, then the elements of one column are obtainable by multiplying all elements of the other by some number. After factoring this number out, we get a determinant with two identical columns; by the property 3, the value of this determinant is zero.

For example,

$$\begin{vmatrix} 8 & 4 & 7 \\ 10 & 5 & 9 \\ 6 & 3 & 11 \end{vmatrix} = 2 \begin{vmatrix} 4 & 4 & 7 \\ 5 & 5 & 9 \\ 3 & 3 & 11 \end{vmatrix} = 0.$$

Property 7. *If each element in the n th column (or the n th row, of a determinant is the sum of two terms, the determinant may be expressed as the sum of two determinants, of which one has in its n th column (or row) the first of the above-mentioned terms, while the other determinant has the second terms; the elements of the remaining columns (or rows) are the same for all the three determinants.*

Thus, for example,

$$\begin{vmatrix} a'_1 + a''_1 & b_1 & c_1 \\ a'_2 + a''_2 & b_2 & c_2 \\ a'_3 + a''_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a'_1 & b_1 & c_1 \\ a'_2 & b_2 & c_2 \\ a'_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} a''_1 & b_1 & c_1 \\ a''_2 & b_2 & c_2 \\ a''_3 & b_3 & c_3 \end{vmatrix}.$$

To prove this relation, we have merely to expand the determinants on both sides of it, applying the rule of triangles, and to compare the results.

Property 8. *If to the elements of a column (or row) of a determinant are added the corresponding elements of another column (or row), multiplied by any one number, the value of the determinant remains unchanged.*

This property follows from the properties 7 and 6, as will be made clear by an example. Let the elements of the second column, each multiplied by some number k , be added to the elements of the first column. Then, by the property 7, we have

$$\begin{vmatrix} a_1 + kb_1 & b_1 & c_1 \\ a_2 + kb_2 & b_2 & c_2 \\ a_3 + kb_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} kb_1 & b_1 & c_1 \\ kb_2 & b_2 & c_2 \\ kb_3 & b_3 & c_3 \end{vmatrix}.$$

The second of the resulting determinants has two proportional columns. Hence, by the property 6, its value is zero, so that we obtain the relation

$$\begin{vmatrix} a_1 + kb_1 & b_1 & c_1 \\ a_2 + kb_2 & b_2 & c_2 \\ a_3 + kb_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix},$$

which expresses the property 8 in the present case.

Further properties of determinants are connected with the concept of cofactors and minors.

§ 4. Cofactors and Minors

11. Consider the determinant

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \quad (1)$$

By definition (see Art. 8),

$$\Delta = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3 - b_1a_2c_3 - a_1c_2b_3. \quad (2)$$

Let us enclose in parentheses all terms containing any one element of this determinant, and factor the element out; the quantity remaining within the parentheses is called *the cofactor* of that element. We shall denote the cofactor of an element by the capital letter and subscript corresponding to the letter and subscript of the element; for example, the cofactor of the element a_1 will be denoted by A_1 , the cofactor of b_1 by B_1 , etc.

Property 9. *A determinant is equal to the sum of the products of the elements in any column (or row) by their cofactors.*

In other words, we have the following relations:

$$\Delta = a_1A_1 + a_2A_2 + a_3A_3, \quad \Delta = a_1A_1 + b_1B_1 + c_1C_1, \quad (3)$$

$$\Delta = b_1B_1 + b_2B_2 + b_3B_3, \quad \Delta = a_2A_2 + b_2B_2 + c_2C_2, \quad (4)$$

$$\Delta = c_1C_1 + c_2C_2 + c_3C_3, \quad \Delta = a_3A_3 + b_3B_3 + c_3C_3. \quad (5)$$

To prove, say, the first of these relations, it is sufficient to rewrite the right-hand member of (2) as

$$\Delta = a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1),$$

where the quantities in the parentheses are the respective cofactors of the elements a_1, a_2, a_3 , that is,

$$b_2c_3 - b_3c_2 = A_1; \quad b_3c_1 - b_1c_3 = A_2; \quad b_1c_2 - b_2c_1 = A_3.$$

Hence the above equality becomes

$$\Delta = a_1A_1 + a_2A_2 + a_3A_3,$$

as was to be shown. The remaining relations (3 through 5) are proved in a similar fashion. The representation of a determinant according to any one of the formulas (3 through 5) is referred to as the expansion of the determinant in terms of the elements of a column or row (thus, the first of those formulas gives the expansion in terms of the elements of the first column, etc.).

12. The determinant obtained from the given determinant by deleting the row and the column, in the intersection of which a particular element lies, is called *the minor* of that element. For example, the determinant $\begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix}$ is the minor of the element a_1 of Δ ; the determinant $\begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix}$ is the minor of b_1 , and so on.

The cofactor of any element of a determinant is equal to the minor of that element taken with its sign unchanged if the sum of the position numbers of the row and column in which the element lies is even, or taken with opposite sign if this sum is odd. To verify this statement, the reader should compare the cofactors of all elements of a determinant with their respective minors.

The fact just stated greatly facilitates the use of formulas (3 through 5), since it enables us to write at once the cofactors of the elements of a determinant (by mere inspection of this determinant). The following diagram will also be of help:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix};$$

the plus signs mark here the positions of those elements whose cofactors are equal to the minors taken with their signs unchanged.

Example. Evaluate the determinant

$$\Delta = \begin{vmatrix} 2 & 4 & 6 \\ 5 & 12 & 19 \\ 3 & 9 & 17 \end{vmatrix}$$

by expanding it in terms of the elements of the first row.

Solution.

$$\Delta = 2 \begin{vmatrix} 12 & 19 \\ 9 & 17 \end{vmatrix} - 4 \begin{vmatrix} 5 & 19 \\ 3 & 17 \end{vmatrix} + 6 \begin{vmatrix} 5 & 12 \\ 3 & 9 \end{vmatrix} = 8.$$

Note. The evaluation of a determinant by expanding it in terms of the elements of a column or row can be simplified if the determinant is first transformed on the basis of the property 8. Namely, if we multiply the elements of a column (or row) by any factor and add them to the corresponding elements of another column (or row), we obtain a new determinant equal to the original one; by a judicious choice of the factor, one of the elements of the new determinant can be made zero. Repeating this procedure once again, we can obtain a determinant (equal to the given one) with two zero elements in a particular row or column. When evaluating the last determinant by expanding it in terms of the elements of this particular row (or column), we shall have to compute only one minor, since the other two minors will be multiplied by elements equal to zero. Thus, for instance, to evaluate the determinant Δ of the above example, it should first be transformed as follows: multiply the elements of the first column by (-2) and add them to the elements of the second column; next, multiply the elements of the first column by (-3) and add them to the elements of the third column; this will give

$$\Delta = \begin{vmatrix} 2 & 0 & 0 \\ 5 & 2 & 4 \\ 3 & 3 & 8 \end{vmatrix}.$$

Expanding this determinant in terms of the first row, we find

$$\Delta = 2 \cdot \begin{vmatrix} 2 & 4 \\ 3 & 8 \end{vmatrix} - 0 \cdot \begin{vmatrix} 5 & 4 \\ 3 & 8 \end{vmatrix} + 0 \cdot \begin{vmatrix} 5 & 2 \\ 3 & 3 \end{vmatrix} = 8.$$

13. The propositions contained in this article are of importance in solving and investigating a system of first-degree equations in three unknowns *).

Consider the determinant

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \quad (1)$$

*) Analogous propositions referring to determinants of higher orders are used in solving and investigating a system of first-degree equations in any number of unknowns.

Expand it in terms of the elements of a row or column, say of the first column:

$$\Delta = a_1A_1 + a_2A_2 + a_3A_3. \quad (6)$$

Replace the numbers a_1, a_2, a_3 in the right member of (6) by any numbers h_1, h_2, h_3 ; then the right member of (6) will be the expansion, in terms of the first column, of the determinant obtained by replacing the elements of the first column of Δ by the numbers h_1, h_2, h_3 :

$$h_1A_1 + h_2A_2 + h_3A_3 = \begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix}. \quad (7)$$

Now let the elements of the second or the third column of the given determinant be chosen as h_1, h_2, h_3 (that is, let $h_1 = b_1, h_2 = b_2, h_3 = b_3$; or let $h_1 = c_1, h_2 = c_2, h_3 = c_3$). In each case the determinant (7) will have two identical columns and will, therefore, be zero; we thus obtain

$$b_1A_1 + b_2A_2 + b_3A_3 = 0, \quad (8)$$

and

$$c_1A_1 + c_2A_2 + c_3A_3 = 0. \quad (9)$$

Expanding the determinant Δ in terms of its second column and using the same procedure as before, we obtain

$$a_1B_1 + a_2B_2 + a_3B_3 = 0, \quad (10)$$

$$c_1B_1 + c_2B_2 + c_3B_3 = 0. \quad (11)$$

Expanding Δ in terms of its third column gives the relations

$$a_1C_1 + a_2C_2 + a_3C_3 = 0, \quad (12)$$

$$b_1C_1 + b_2C_2 + b_3C_3 = 0. \quad (13)$$

Six similar relations can also be written for the rows of the determinant.

According to these results, we may formulate the following property of determinants:

Property 10. *The sum of the products of the elements in any column (or row) by the cofactors of the corresponding elements in another column (or row) is zero.*

§ 5. Solution and Analysis of a System of Three First-degree Equations in Three Unknowns

14. Consider the system of three equations

$$\left. \begin{aligned} a_1x + b_1y + c_1z &= h_1, \\ a_2x + b_2y + c_2z &= h_2, \\ a_3x + b_3y + c_3z &= h_3 \end{aligned} \right\} \quad (1)$$

in the unknowns x, y, z (the coefficients a_1, b_1, \dots, c_3 and the constant terms h_1, h_2, h_3 are assumed to be known).

The three numbers x_0, y_0, z_0 are said to be a solution of the system (1) if these numbers satisfy the equations of the system (1), that is, if each of equations (1) becomes an arithmetical identity after substituting the numbers x_0, y_0 and z_0 for x, y and z , respectively. Let us find all solutions of the system (1); at the same time, we shall analyse the system, namely, we shall determine the cases where the system (1) has only one solution, more than one solution, and no solution at all.

The determinant

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad (2)$$

formed from the coefficients of the unknowns is called *the determinant of the system*; it plays a basic part in the discussion that follows.

Let the symbols A_1, A_2, \dots denote, as before, the cofactors of the elements a_1, a_2, \dots of the determinant Δ . Multiplying the first equation of (1) throughout by A_1 , the second equation by A_2 , the third by A_3 , and adding them, term by term, we obtain

$$(a_1A_1 + a_2A_2 + a_3A_3)x + (b_1A_1 + b_2A_2 + b_3A_3)y + (c_1A_1 + c_2A_2 + c_3A_3)z = (h_1A_1 + h_2A_2 + h_3A_3).$$

Hence, by virtue of the properties 9 and 10 (see the first of relations (3) in Art. 11, and also relations (8), (9) in Art. 13), we have

$$\Delta \cdot x = h_1A_1 + h_2A_2 + h_3A_3. \quad (3)$$

In like manner, we find

$$\Delta \cdot y = h_1B_1 + h_2B_2 + h_3B_3, \quad (4)$$

$$\Delta \cdot z = h_1C_1 + h_2C_2 + h_3C_3. \quad (5)$$

Denote the right members of (3), (4) and (5) by the symbols Δ_x , Δ_y , Δ_z , respectively. Then equations (3), (4), (5) will take the form

$$\Delta \cdot x = \Delta_x, \quad \Delta \cdot y = \Delta_y, \quad \Delta \cdot z = \Delta_z, \quad (6)$$

where, according to Art. 13 (see, for example, formula (7) of Art. 13),

$$\Delta_x = \begin{vmatrix} h_1 & b_1 & c_1 \\ h_2 & b_2 & c_2 \\ h_3 & b_3 & c_3 \end{vmatrix}, \quad \Delta_y = \begin{vmatrix} a_1 & h_1 & c_1 \\ a_2 & h_2 & c_2 \\ a_3 & h_3 & c_3 \end{vmatrix}, \quad \Delta_z = \begin{vmatrix} a_1 & b_1 & h_1 \\ a_2 & b_2 & h_2 \\ a_3 & b_3 & h_3 \end{vmatrix}. \quad (7)$$

It will be helpful to note that the determinants Δ_x , Δ_y , Δ_z are obtained from Δ by replacing its first, second and third column, respectively, by the column of the constant terms of the given system.

Suppose that $\Delta \neq 0$; from equations (6), we then find

$$x = \frac{\Delta_x}{\Delta}, \quad y = \frac{\Delta_y}{\Delta}, \quad z = \frac{\Delta_z}{\Delta}. \quad (8)$$

These formulas obviously give the solution of the derived system consisting of equations (6). They also give the solution of the original system (1). To prove this, expressions (8) must be substituted for x , y , z in the equations of the system (1) and shown to satisfy each of equations (1). Making these substitutions in the first equation, we obtain

$$\begin{aligned} & a_1 \frac{\Delta_x}{\Delta} + b_1 \frac{\Delta_y}{\Delta} + c_1 \frac{\Delta_z}{\Delta} = \\ &= \frac{1}{\Delta} a_1 (h_1 A_1 + h_2 A_2 + h_3 A_3) + \frac{1}{\Delta} b_1 (h_1 B_1 + h_2 B_2 + h_3 B_3) + \\ &+ \frac{1}{\Delta} c_1 (h_1 C_1 + h_2 C_2 + h_3 C_3) = \frac{1}{\Delta} h_1 (a_1 A_1 + b_1 B_1 + c_1 C_1) + \\ &+ \frac{1}{\Delta} h_2 (a_1 A_2 + b_1 B_2 + c_1 C_2) + \frac{1}{\Delta} h_3 (a_1 A_3 + b_1 B_3 + c_1 C_3). \end{aligned}$$

Now, by the property 9 of determinants,

$$a_1 A_1 + b_1 B_1 + c_1 C_1 = \Delta,$$

and, by the property 10,

$$a_1 A_2 + b_1 B_2 + c_1 C_2 = 0,$$

$$a_1 A_3 + b_1 B_3 + c_1 C_3 = 0.$$

Thus,

$$a_1 \frac{\Delta_x}{\Delta} + b_1 \frac{\Delta_y}{\Delta} + c_1 \frac{\Delta_z}{\Delta} = h_1,$$

that is, the numbers x, y, z , determined by formulas (8), satisfy the first equation of the given system; in exactly the same way, they can be shown to satisfy the other two equations.

From all this, the following conclusion may be drawn: *If $\Delta \neq 0$, the system (1) has a unique solution determined by formulas (8).*

Example. Find all solutions of the system

$$\left. \begin{aligned} x + 2y + z &= 4, \\ 3x - 5y + 3z &= 1, \\ 2x + 7y - z &= 8. \end{aligned} \right\}$$

Solution. Let us evaluate the determinant of the system.

$$\Delta = \begin{vmatrix} 1 & 2 & 1 \\ 3 & -5 & 3 \\ 2 & 7 & -1 \end{vmatrix} = 33.$$

Since $\Delta \neq 0$, the system has a unique solution given by formulas (8). From formulas (7), we have

$$\Delta_x = \begin{vmatrix} 4 & 2 & 1 \\ 1 & -5 & 3 \\ 8 & 7 & -1 \end{vmatrix} = 33, \quad \Delta_y = \begin{vmatrix} 1 & 4 & 1 \\ 3 & 1 & 3 \\ 2 & 8 & -1 \end{vmatrix} = 33,$$

$$\Delta_z = \begin{vmatrix} 1 & 2 & 4 \\ 3 & -5 & 1 \\ 2 & 7 & 8 \end{vmatrix} = 33.$$

Hence $x = 1, y = 1, z = 1$.

15. Suppose now that the determinant of the system (1) is zero: $\Delta = 0$.

If $\Delta = 0$ and at least one of the determinants $\Delta_x, \Delta_y, \Delta_z$ is different from zero, the system (1) has no solution (the equations of the system are said to be inconsistent).

For, if $\Delta = 0$, but at least one of the determinants $\Delta_x, \Delta_y, \Delta_z$ is not equal to zero, then at least one of equations (6) is impossible, which means that the system (6) has no solutions.

But then the system (1) does not possess any solutions either, since the system (6) has been derived from the system (1), and hence each solution of (1), if any such solution should exist, would also be a solution of the system (6). For example, the system

$$\left. \begin{aligned} x + y + z &= 2, \\ 3x + 2y + 2z &= 1, \\ 4x + 3y + 3z &= 4 \end{aligned} \right\}$$

has no solutions, since $\Delta = 0$ and $\Delta_y = 1 \neq 0$. The fact that these equations are inconsistent can also be verified directly; in fact, adding the first two equations, term by term, and subtracting the result from the third equation, we obtain $0 = 1$, that is, an incorrect equation.

There remains for consideration the case where $\Delta = 0$ and also $\Delta_x = 0$, $\Delta_y = 0$, $\Delta_z = 0$. Before taking up this case, we shall first discuss the so-called homogeneous systems.

16. A *homogeneous* system of three first-degree equations in three unknowns is a system of the form

$$\left. \begin{aligned} a_1x + b_1y + c_1z &= 0, \\ a_2x + b_2y + c_2z &= 0, \\ a_3x + b_3y + c_3z &= 0, \end{aligned} \right\} \quad (9)$$

that is, a system of equations whose constant terms are all equal to zero. Clearly, such a system always possesses the solution $x = 0$, $y = 0$, $z = 0$, which is called the *zero* solution. If $\Delta \neq 0$, this solution is unique.

We shall now prove that, if $\Delta = 0$, the homogeneous system (9) has an infinite number of non-zero solutions. (In this event, one of its equations is a consequence of the other two equations, or else two of its equations are consequences of the third.)

Let us first carry out the proof for the case where at least one of the minors of the determinant Δ is different from zero; let, for instance,

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0.$$

Under this condition, the first two equations of the system (9) have an infinite number of simultaneous non-zero solutions determined by the formulas

$$x = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} t, \quad y = - \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} t, \quad z = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} t, \quad (10)$$

for any value of t (see formulas (8) of Art. 6). It can easily be verified that, if $\Delta = 0$, all these numbers also satisfy the third equation of the system (9). For, substituting them for the unknowns in the left-hand member of the third equation of (9), we find

$$a_3x + b_3y + c_3z = \left(a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} + c_3 \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \right) t = \Delta \cdot t.$$

As a result of the substitution, we get zero since, by hypothesis, $\Delta = 0$. Thus, formulas (10) determine, for any value of t , the solutions of the system (9), which will be non-zero solutions when $t \neq 0$. In the case under discussion, the system consists, essentially, of two equations only, the third equation being a consequence of the first two.

Suppose now that the minors of the determinant Δ are all equal to zero; then any two equations of the system (9) have proportional coefficients, so that (no matter which two equations of the system are chosen) one of them can be obtained by multiplying all terms of the other by some common factor (in this connection, see Art. 7). Hence the system (9) will consist, essentially, of only one equation, the other two equations being its consequences. Such a system obviously has an infinite number of non-zero solutions (since we can assign any numerical values to two unknowns and then determine the third unknown from the only essential equation of the system). This completes the proof. The result can be formulated as follows:

The homogeneous system (9) has non-zero solutions if, and only if, $\Delta = 0$.

Example 1. The system

$$\left. \begin{aligned} x + 2y + z &= 0, \\ 3x - 5y + 3z &= 0, \\ 2x + 7y - z &= 0 \end{aligned} \right\}$$

has the zero solution only, since

$$\Delta = 33 \neq 0.$$

Example 2. The system

$$\left. \begin{aligned} x + y + z &= 0, \\ 2x + 3y + 2z &= 0, \\ 4x + 5y + 4z &= 0 \end{aligned} \right\}$$

has infinitely many non-zero solutions, since

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 4 & 5 & 4 \end{vmatrix} = 0.$$

All the solutions are given by formulas (10), according to which

$$x = -t, \quad y = 0, \quad z = t,$$

for any value of t .

Example 3. The system

$$\left. \begin{aligned} x + y + z &= 0, \\ 2x + 2y + 2z &= 0, \\ 3x + 3y + 3z &= 0 \end{aligned} \right\}$$

also has an infinite number of non-zero solutions, since $\Delta = 0$. In the present case, the minors of the determinant Δ are all equal to zero and the system reduces to the single equation $x + y + z = 0$. Any solution of the system consists of three numbers x, y, z , where x, y are chosen at will and $z = -x - y$.

17. Let us return to our arbitrary non-homogeneous system

$$\left. \begin{aligned} a_1x + b_1y + c_1z &= h_1, \\ a_2x + b_2y + c_2z &= h_2, \\ a_3x + b_3y + c_3z &= h_3. \end{aligned} \right\} \quad (1)$$

We shall now prove that, if $\Delta = 0$ and the system (1) has at least one solution, then the system has an infinite number of different solutions.

Let the numbers x_0, y_0, z_0 constitute a solution of the system (1); substituting x_0, y_0, z_0 for the unknowns in equations (1), we obtain the arithmetical identities

$$\left. \begin{aligned} a_1x_0 + b_1y_0 + c_1z_0 &= h_1, \\ a_2x_0 + b_2y_0 + c_2z_0 &= h_2, \\ a_3x_0 + b_3y_0 + c_3z_0 &= h_3. \end{aligned} \right\} \quad (11)$$

Subtracting identities (11), term by term, from the corresponding equations (1), that is, subtracting the first identity (11) from the first equation of (1), the second identity from the second equation, and the third identity from the third equation, we have

$$\left. \begin{aligned} a_1(x - x_0) + b_1(y - y_0) + c_1(z - z_0) &= 0, \\ a_2(x - x_0) + b_2(y - y_0) + c_2(z - z_0) &= 0, \\ a_3(x - x_0) + b_3(y - y_0) + c_3(z - z_0) &= 0. \end{aligned} \right\} \quad (12)$$

Introducing the notation

$$x - x_0 = u, \quad y - y_0 = v, \quad z - z_0 = w, \quad (13)$$

we can rewrite relations (12) as

$$\left. \begin{aligned} a_1u + b_1v + c_1w &= 0, \\ a_2u + b_2v + c_2w &= 0, \\ a_3u + b_3v + c_3w &= 0. \end{aligned} \right\} \quad (14)$$

This is a homogeneous system of three first-degree equations in the unknowns u, v, w , with the same coefficients of the unknowns as in the original system (1). The system (14) is referred to as the homogeneous system corresponding to the given non-homogeneous system (1).

Since, by hypothesis, $\Delta = 0$, it follows from Art. 16 that the homogeneous system (14) has infinitely many different solutions. Hence the given system (1) also has infinitely many different solutions; namely, to each solution u, v, w of the system (14) there corresponds—in consequence of relation (13)—a solution

$$x = x_0 + u,$$

$$y = y_0 + v,$$

$$z = z_0 + w$$

of the system (1). The proof is thus complete.

Accordingly, we can immediately formulate the following proposition:

*If $\Delta = 0$ and also $\Delta_x = \Delta_y = \Delta_z = 0$, the system (1) has either no solution, or infinitely many solutions (in the latter case, at least one equation of the system is a consequence of the other equations; such a system is called an *indeterminate system*).*

Example 1. The system

$$\left. \begin{aligned} x + y + z &= 1, \\ 2x + 2y + 2z &= 3, \\ 3x + 3y + 3z &= 4 \end{aligned} \right\}$$

(satisfying the conditions $\Delta = 0$, $\Delta_x = 0$, $\Delta_y = 0$, $\Delta_z = 0$) has no solutions.

In fact, even the first two equations of the system are inconsistent since; multiplying the first of them by 2 and then subtracting it, term by term, from the second equation, we obtain the impossible equality $0 = 1$.

Example 1. The system

$$\left. \begin{aligned} 3x + y - z &= 1, \\ 5x + 2y + 3z &= 2, \\ 8x + 3y + 2z &= 3 \end{aligned} \right\}$$

(satisfying the conditions $\Delta = 0$, $\Delta_x = 0$, $\Delta_y = 0$, $\Delta_z = 0$) has infinitely many solutions. For, the third equation of the system is a consequence of the first two; namely, it can be obtained by adding them, term by term. Thus, the given system consists, essentially, of only two equations,

$$\left. \begin{aligned} 3x + y - z &= 1, \\ 5x + 2y + 3z &= 2. \end{aligned} \right\} \quad (*)$$

To find all simultaneous solutions of equations (*), rewrite the system as

$$\left. \begin{aligned} 3x + y &= 1 + z, \\ 5x + 2y &= 2 - 3z, \end{aligned} \right\}$$

and suppose that some value is assigned here to the unknown z . Applying formulas (8) of Art. 2, we get

$$x = \frac{\begin{vmatrix} 1+z & 1 \\ 2-3z & 2 \end{vmatrix}}{\begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix}} = 5z, \quad y = \frac{\begin{vmatrix} 3 & 1+z \\ 5 & 2-3z \end{vmatrix}}{\begin{vmatrix} 3 & 1 \\ 5 & 2 \end{vmatrix}} = 1 - 14z.$$

The numbers x , y , together with the number z , constitute a solution of the given system; the system has infinitely many solutions since the unknown z may be assigned any value at will.

§ 6. The Concept of a Determinant of Any Order

18. The general problem of solving and analysing systems of first-degree equations in many unknowns (and a large number of other computational problems of mathematics) has to do with determinants of the n th order ($n = 2, 3, 4, \dots$). The theory of determinants of any order is built along the same general lines as the above-presented theory of determinants of the third order; a detailed development of the theory of determinants of any order would, however, involve a number of new auxiliary propositions and would thus offer certain difficulties. This theory, as well as the theory of systems of first-degree equations in many unknowns, is given in textbooks on higher algebra.

We shall confine ourselves to stating the following:

1. A determinant of order n is associated with a square array of numbers (elements of the determinant) arranged in n rows and n columns; the notation for a determinant of order n is analogous to that used for determinants of the 2nd and the 3rd order.

2. The minor of an element in a determinant of order n is defined as the determinant of order $n - 1$ obtained from the given determinant by striking out the row and column in which that element appears.

3. The *cofactor* of an element of a determinant is the minor of that element, taken with its sign unchanged if the sum of the position numbers of the row and column in which the element lies is even, or taken with opposite sign if this sum is odd.

4. A determinant is equal to the sum of the products of the elements in any column (or row) by their cofactors. This reduces the evaluation of a determinant of order n to the evaluation of n determinants of order $n - 1$.

5. All these properties of determinants hold for determinants of any order.

Example. Evaluate the determinant

$$\Delta = \begin{vmatrix} 2 & 4 & 4 & 6 \\ 4 & 2 & 5 & 7 \\ 3 & 2 & 8 & 5 \\ 2 & 8 & 7 & 3 \end{vmatrix}.$$

Solution. Expanding this determinant in terms of the upper row, that is, writing out the determinant as the sum of the products of the elements in the upper row by their cofactors, we find

$$\Delta = 2 \begin{vmatrix} 2 & 5 & 7 \\ 2 & 8 & 5 \\ 8 & 7 & 3 \end{vmatrix} - 4 \begin{vmatrix} 4 & 5 & 7 \\ 3 & 8 & 5 \\ 2 & 7 & 3 \end{vmatrix} + 4 \begin{vmatrix} 4 & 2 & 7 \\ 3 & 2 & 5 \\ 2 & 8 & 3 \end{vmatrix} - 6 \begin{vmatrix} 4 & 2 & 5 \\ 3 & 2 & 8 \\ 2 & 8 & 7 \end{vmatrix} = 296.$$

Note. The evaluation of the determinant can be simplified by using first the property 8 (see Art. 10 and the note at the end of Art. 12).

INDEX

A

- Abscissa, 17, 18, 132
- Amplitude, 21
- Angle, between two lines, 59
 - between two vectors, 161
 - of inclination, 55, 136
 - polar, 21
 - of a segment, 25
- Applicate, 132
- Area of a triangle, 28
- Associative property, of multiplication of a vector by numbers, 147
 - of the scalar product, 158
 - of the vector product, 166
 - of vector addition, 145
- Asymptote, 94
- Axes, coordinate, 17, 131
 - of an ellipse, 81
 - of a hyperbola, 95
- Axis, of a parabola, 106
 - polar, 20
 - semi-major, 81
 - semi-minor, 81

B

- Basis, 154
- Branch of a hyperbola, 88

C

- Canonical equation, of an ellipse, 76
 - of a hyperbola, 87
 - of a parabola, 103
 - of a second-order curve, 117, 120
- Cartesian coordinates, left-handed system, 40
 - oblique, 90
 - rectangular, 16, 132
 - right-handed system, 40

- Centre, of an ellipse, 81
 - of gravity, 34
 - of a hyperbola, 95
 - of a second-order curve, 119
- Circle, equations of, 44, 48, 49
 - relation of, to the ellipse, 81
- Cofactor, 236, 247
- Commutative property, not valid for
 - the vector product, 165
 - of the scalar product, 158
 - of vector addition, 144
- Components of a vector, 154
- Cone, circular, 214
 - imaginary, 215
 - quadric, 214
- Conic sections, 115
- Conjugate hyperbolas, 97
- Construction of an ellipse, 76, 83
- Coordinates, current, 43, 179
 - oblique cartesian, 20
 - of a vector, 139
 - on a line, 14
 - polar, 20
 - rectangular cartesian, 16, 131
 - relations between rectangular and polar, 21
- Curve, algebraic, 53
 - equation of, 42
- Cylinder, circular, section of, 85
 - degenerate, 220
 - imaginary, 220
 - projecting, 182
 - quadric, types of, 219

D

- Departure of a point, from a line, 70
 - from a plane 191
- Determinants, of the second order, 225
 - of the third order, 232
 - properties of, 233, 247

Diameters, conjugate, 113
 of a conic, 109
 Directed segment, 11
 Direction cosines, 141, 199
 Direction parameters, 199
 Direction vector, 198
 Directrix, of an ellipse, 99
 of a hyperbola, 100
 of a parabola, 103
 Distance, between two points, 25, 142
 of a point from a line, 70
 of a point from a plane, 191
 Distributive property, of multiplication
 of vectors by numbers, 147
 of the scalar product, 158
 of the vector product, 167

E

Eccentricity, of an ellipse, 81
 of a hyperbola, 98
 of a parabola, 103
 Ellipse, 75
 degenerate, 122
 imaginary, 122
 optical property of, 114
 Ellipsoid, 207
 imaginary, 209
 of revolution, oblate, 209
 prolate, 208

F

Focal radii, of an ellipse, 75, 82
 of a hyperbola, 88, 98
 Focus, of an ellipse, 75
 of a hyperbola, 87
 of a parabola, 103
 Function, linear, 58
 quadratic, 127

G

General equation, of a plane, 186
 of a second-order curve, 117
 of a straight line, 62
 Graph of a function, 42

H

Half-planes, 18
 Half-spaces, 133
 Homogeneous equations, 213
 system of, 229, 242

Hyperbola, 87
 degenerate, 122
 equilateral, 97
 optical property of, 114
 Hyperboloid, of one sheet, 209
 of two sheets, 209

I

Identity, fundamental, 12
 Incomplete equations, of planes, 188
 of straight lines, 63
 Intercept, 57, 65
 Intercept equation, of a plane, 189
 of a straight line, 63
 Intersection, of curves, 50
 of lines, 66
 of surfaces, 180

M

Midpoint of a segment, 33, 143
 Minor, 237, 247
 Modulus, 12
 of a vector, 136, 140
 Motion, equations of, 51, 200

N

Negative of a vector, 148
 Normal, to a line, 68
 to a plane, 190
 Normal equation, of a plane, 190
 of a straight line, 68
 Number axis, 16

O

Octants, 134
 Ordinate, 17, 132
 Origin of coordinates, 17, 131

P

Parabola, 103
 degenerate, 125
 optical property of, 114
 Paraboloid, elliptic, 215
 hyperbolic, 216
 Parallel lines, 60, 67, 203
 Parallel planes, 202
 Parallelogram rule, 144
 Parameter, 41
 focal, 108
 of a parabola, 103

Parametric equations, 51
 of an ellipse, 83
 of spirals, 52
 of a straight line, 200

Pencil, of lines, 71
 of planes, 197

Perpendicular lines, 60, 203

Perpendicular planes, 202

Plane, 185

through a line, 195
 through a point, 186
 through three points, 201

Polar coordinates, 20

Polar equation, of a circle, 45

of a conic, 107
 of a straight line, 50

Pole, 20

Principal value, 21

Projection, of a segment, 23
 of a vector, 136, 148, 155, 162

Product, of a vector by a number, 143
 scalar, 158
 triple scalar, 172
 vector, 163

Q

Quadrants, 19

Quadric surfaces, 207

R

Radius vector, 24, 140

Ratio of division, 31, 142

Rectangle, fundamental, 96

Rectangular coordinates, in the plane, 16

in space, 132

Removal of xy -term, 118, 120

Rotation of coordinate axes, 36

Ruled surfaces, 220

S

Scalar, 134

Shukhov structures, 224

Slope, 55

Slope-intercept equation of a straight line, 56

Spiral, hyperbolic, 46

logarithmic, 47

of Archimedes, 45

Straight line, 55

in space, 194

through a point, 58, 199

through two points, 58, 201

Subtraction of vectors, 147

Sum of vectors, 143, 145

Surface, algebraic, 184

equation of, 178

quadric, 207

System, of two first-degree equations, 224

of three first-degree equations, 240

T

Transformation, of coordinates, 21, 35, 36, 37

of equations, 116

Translation of coordinate axes, 35

Triad of vectors, 173

V

Value of a segment, 11

Vertex, of a cone, 214

of an ellipse, 81

of a hyperbola, 96

of a parabola, 106

of a paraboloid, 216, 218

Vector, 134

direction, 198

free, 136

geometric, 134

normal, 186

Vectors, collinear, 135

coplanar, 173, 177

equal, 135

perpendicular, 161

Volume of a tetrahedron, 177

To the Reader

Peace Publishers would be glad to have your opinion of the translation and the design of this book.

Please send your suggestions to 2, Pervy Rizhsky Pereulok, Moscow, U. S. S. R.



Printed in the Union of Soviet Socialist Republics

